OAK RIDGE NATIONAL LABORATORY

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

ORNL/TM-2009/147

NYISO Industrial Load Response Opportunities: Resource and Market Assessment—Task 2 Final Report

October 2009

Prepared by

B. Kirby
M. Starke
S. Adhikari

Oak Ridge National Laboratory,
Oak Ridge, Tennessee

UT-BATTELLE

ORNL-27 (4-00)

ORNL/TM-2009/147

NYISO Industrial Load Response Opportunities: Resource and Market Assessment—Task 2 Final Report

October 2009

Prepared for

New York State Energy Research and Development Authority
Albany, NY
Joseph H. Sayer, Senior Project Manager

U.S. Department of Energy Office of Electricity Delivery and Energy Reliability
Washington, DC
Philip N. Overholt, Transmission Reliability Program Manager

Alcoa Inc.
New York, NY
Michael W. Caufield, Energy Regulatory Affairs Specialist

by

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6283
managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-76RL01830

ABSTRACT

This report examines the ability of responsive loads to reduce energy costs and increase power system reliability by selling ancillary services (regulation, contingency reserves, and dynamic reactive support) to the power system while purchasing power to perform the basic functions the loads were designed to do. A basic understanding of power system ancillary services is provided including definitions and characteristics of each service. Preliminary examination of ancillary service prices in NY, Texas, California, and New England indicate that regulation prices are particularly attractive, and have been for years. Studying the hourly energy and ancillary service market prices provides a sound basis for determining the genuine value of response to the power system and the revenue responsive loads can likely capture. Load response is examined as it relates to power system needs in general and as it relates to the specific requirements of the power system operated by the New York Independent System Operator. The various current opportunities for load response within the New York Independent System Operator market structure are discussed. A modeling effort to determine the optimal mix of energy and ancillary service response capabilities for individual loads is outlined.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ACE	area control error
AGC	automatic generation control
CAISO	California Independent System Operator
CPS	control performance standard (1&2)
DADRP	day-ahead demand response program
DAMAP	day-ahead margin assurance payment
DCS	disturbance control standard
DSASP	demand side ancillary service program
D-VAR	American Superconductor’s dynamic VAR reactive compensation systems
DRP	demand response program
EDRP	emergency demand response program
ERCOT	Electric Reliability Council of Texas
FERC	Federal Energy Regulatory Commission
FRR	frequency responsive reserve
ICAP/SCR	installed capacity / special case resources
IPP	independent power producer
ISO	independent system operator
ISO-NE	Independent System Operator–New England
Kvar	thousands of volt-amps reactive
LGIA	Large Generator Interconnection Agreement
LGIP	Large Generator Interconnection Procedures
LMP	locational marginal price
LBMP	locational based marginal pricing
LSE	load serving entity
MDSP	meter data service providers
MISO	Midwest Independent System Operator
MVA	millions of volt-amps
Mvar	millions of volt-amps reactive
MW	megawatt
MWh	megawatt-hour of energy
$/MW-hr	price unit for reserves, one MW of *capacity* (not energy) for one hour
NERC	North American Electric Reliability Council
NYISO	New York Independent System Operator
RIP	responsible interface parties
RMR	reliability must run
RTO	regional transmission organization
SCADA	supervisory control and data acquisition
SCR	special case resource
UCAP	unforced capacity
UPS	uninterruptible power supply
WECC	Western Electricity Coordinating Council

SUMMARY

This report examines the ability of responsive loads to reduce energy costs by responding to power system needs. Specifically, this report examines the technical and economic potential for loads to sell regulation, contingency reserves, and dynamic reactive support to the power system while purchasing power to perform the basic functions the loads were designed to do. Interactions with the real-time electric power markets are examined in order to optimize overall electric power procurement.

Ancillary services are procured through hourly competitive markets by the major independent system operators (ISOs). Ancillary service prices are driven by the opportunity costs of the generators who currently dominate the ancillary service supply markets. Consequently, ancillary service prices are intertwined with real-time energy prices, vary from ISO to ISO along with the generation technology mix, and vary with time-of-day, day-of-week, and season. Historical hourly energy and ancillary service price data from these markets form a good basis for estimating the value of providing various types of load response to the power system.

To determine the viability of a resource, it is important to consider the fundamental value of load response that can be provided to each ISO. Furthermore, the specific market and reliability rules that define the services being compensated must also be considered. Market and reliability rules have developed based on the capabilities and interests of the incumbent supply technology (generation). Rules often do not accommodate new technologies until those technologies demonstrate their viability. Discriminatory rules can often be changed if a new technology genuinely meets power system technical reliability requirements and provides fundamental value. The Federal Energy Regulatory Commission (FERC) has provided clear signals that utilizing load response whenever viable should be encouraged. However, it can take months or even years to change these rules.

Monitoring requirements are often the largest obstacle to aggregations of small loads selling ancillary services to the power system; for example, the North American Electric Reliability Council (NERC) reliability rules used to specifically prohibit loads from providing spinning reserve. Although NERC and most Regional Reliability Councils have eliminated that specific restriction, monitoring requirements designed for large resources are still in place. Because they are subject to change, these requirements should be viewed separately from more fundamental perspectives such as the price/value of response. Still, these requirements are important short-term limitations on the type of response that can be sold.

Several ISOs provide capacity payments as well as payments for energy and ancillary services. These payments can amount to a considerable fraction of the total supplier income. They also typically depress the real-time prices for both energy and ancillary services. Unfortunately, capacity payments are typically structured around meeting or reducing peak demand and/or supplying continuous energy response. Capacity qualification rules are also typically detailed and specific to the ISO. They are subject to change since they are not tied tightly to fundamental economic and reliability requirements.

Reactive power, especially dynamic reactive power, is as critical to power system reliability as real power. Responsive loads with solid state front ends have the potential to supply dynamic reactive power to the power system. Still, high reactive impedance within the power system makes reactive power a local issue. Consequently, regional markets are not operated for reactive power, and compensation for the supply of dynamic reactive reserves is typically negotiable. Compensation can be based on the cost of transmission-based, dynamic reactive alternatives such as Static Var Compensators or STATCOMs. Payment can also be based on the cost of generation alternatives such as reliability must-run (RMR) contracts. In some cases, dynamic reactive tariffs exist. Finding specific business opportunities for the supply of dynamic reactive power is significantly more difficult than determining the value of supplying the real-power ancillary services.

Since dynamic reactive power can potentially be supplied by adjustable speed motor drives and solid state power supplies, it may be possible to supply dynamic reactive power 24 hours per day, independently of the load's real power consumption or the supply of other ancillary services. Supplying dynamic reactive power typically requires over-sizing the power electronics so there is a capital cost associated with providing the capability. The dynamic reactive capability requirement of the host utility and the necessary payback options always need to be assessed before installing the capability in locations where it is incrementally profitable.

Data analysis takes several forms. Much can be learned by examining ancillary service price patterns. There are price differences among regions as well as daily and seasonal price patterns. Regulation is always the most expensive ancillary service, followed by spinning reserve, non-synchronous reserve, and supplemental reserve. It can also be useful to model energy consumption and ancillary service provision to determine what response capabilities and operating strategies are the most profitable. As a general rule, response speed is of greater value than response duration.

Sub-hourly energy markets provide another potential opportunity for responsive loads. ISOs clear real-time energy markets every 5 minutes. While the annual average prices in day-ahead-hourly, hour-ahead-hourly, and 5-minute real-time energy markets are typically quite close, the within-hour price spread for 5-minute markets is typically quite large: $90/MWH in New York during 2008. It is not completely clear how a responsive load could exploit this price spread or if it would be profitable to do so, but data are available to facilitate analysis to answer the question.

All power systems require ancillary services. ISOs that operate hourly ancillary service markets provide price transparency, making it easy to see the value of generation and demand response. The services are also valuable in areas without markets, but it is more difficult for alternative suppliers, like demand response, to be allowed to perform and be paid for that performance.

1. INTRODUCTION

Industrial loads have a range of power system market opportunities in New York State. The New York Independent System Operator (NYISO) operates a variety of ancillary service markets that can provide industrial customers with alternative revenue sources. Responsive loads can also participate in capacity and emergency response programs. Five-minute energy markets potentially provide an additional opportunity for loads to benefit financially from their ability to respond to power system needs.

Traditionally, ancillary services (reliability services) have been supplied by generators that respond to power system operator commands. Technically, load response can be as effective as generation response for restoring the generation/load balance at the center of most power system reliability concerns. Loads have demonstrated that they can provide response as fast and reliable as that provided by generation. NYISO markets now largely recognize the value of load response and allow loads to participate equally in most markets.

Power systems require ancillary services to maintain reliability and support their primary function of delivering energy to customers. Ancillary services are principally real-power generator control capacity services the system operator uses over various time frames to maintain the required instantaneous and continuous balance between aggregate generation and load. The NYISO obtains ancillary services by operating hourly markets where suppliers offer response for compensation. Four market-based ancillary services are discussed in this report. Regulation is used continuously to balance generation and load under normal conditions. The remaining three ancillary services only respond intermittently to reliability events but must continuously stand ready to respond: spinning reserve, non-synchronous reserve, and 30-minute reserve. Voltage support is the provision of dynamic reactive power and is discussed as well. Since voltage support is location specific, it is not currently practical to operate hourly markets to obtain the service. Black start is another ancillary service procured by the NYISO but not through hourly markets. Black start is fundamentally a generation service, which limits the ability of loads to supply the service.

Ancillary services are not new. The functions have been provided by vertically integrated utilities since power systems began to be formed a century ago. With restructuring, there is a need to more carefully define, measure, and pay for these services. Advances in communications and controls make it possible for some loads to profitably supply ancillary services to the power system. Transparent markets provide clear economic signals telling loads (and generators) the worth of specific types of response. Loads can then evaluate their ability to profit from providing the desired response. The load can benefit from additional income while the power system and its customers benefit from increased reliability resources and lower costs.

Fundamental power system reliability requirements are quite stable. However, ancillary service and energy market rules are not. This report addresses the fundamental value of load response as well as the current market rules and obstacles to make full use of load flexibility. Understanding the underlying physical power system reliability requirements can help identify when market and reliability rules should be changed to help system operators take advantage of the unique capabilities of responsive load.

This report is organized into a summary and nine sections. Section 1 is this introduction. Section 2 discusses power system reliability needs and the ancillary services themselves, why the power system needs each, and the physical requirements for supplying them. Section 3 provides a brief discussion of the types of demand response that are useful to the power system and limitations on some loads' abilities to provide response. Section 4 discusses the ancillary service and energy price data that are available from NYISO. Section 5 discusses modeling of load response while supplying ancillary services. Modeling can determine how profits can be maximized by selecting how much of each ancillary service to sell during each hour of a study year. Section 6 discusses specific NYISO requirements for supplying ancillary services including metering and monitoring requirements. Section 7 discusses sub-hourly energy price variability and the potential to profitably respond to those sub-hourly price variations. Section 8 discusses opportunities for industrial loads within the current and near-future NYISO market structure. Section 9 provides conclusions and suggestions.

2. ANCILLARY SERVICES FOR POWER SYSTEM RELIABILITY

In order to understand how responsive loads can profit by offering to provide reliability response, one must first understand power system reliability requirements and the ancillary services power systems procure to meet those obligations.

The electric power system has two unique requirements that must be continuously and exactly satisfied to maintain overall system stability and reliability. They are (i) the need to maintain a constant balance between generation and load (there is no storage), and (ii) the need to adjust generation (or load) to manage power flows within the constraints of individual transmission facilities (there is no flow control).[1] These requirements have existed since interconnected power systems started to develop a century ago, and vertically integrated utilities have traditionally maintained this continuous balancing act as a normal part of the electricity business. These two principles lead to four important consequences: (i) prices are inherently volatile, (ii) system operations and transmission are communal and must be regulated, (iii) current operations are often restricted by preparations for the next unlikely event, and (iv) response has value.

With restructuring, the multiple functions that vertically integrated utilities performed as part of their bundled service are being explicitly delineated. FERC, through Order 888, Order 889, Order 2000, and continuing reform effort has defined these as "ancillary services…necessary to support the transmission of electric power from seller to purchaser given the obligations of balancing areas and transmitting utilities within those balancing areas to maintain reliable operations of the interconnected transmission system."

The two basic characteristics and the four consequences underlie the need for and value of capacity and the ancillary services. The basics often get lost in the implemented details of reliability and market rules. Energy is still the basic commodity of interest to electricity users. Everything else simply supports the delivery of energy. In its simplest form, operating an interconnected power system can be reduced to a few tasks:

- Balance aggregate generation to aggregate load.
  - Under normal conditions.
  - Under contingency conditions.[2]
- Maintain voltages throughout the power system.
  - Under normal conditions.
  - Under contingency conditions.

[1] This is not strictly true. Electricity can be stored and flow can be controlled on a small scale and/or at high cost. The behavior of interconnected AC power systems, however, is dominated by the lack of practical storage and limited flow control. These two characteristics differentiate electricity from communications systems (telephone, cell, radio, and internet); pipe systems (gas, water, oil, etc.); and transportation (air, rail, road, and sea). The result is a very different set of control requirements and ancillary services.

[2] A contingency is the sudden, unexpected loss of a generator or transmission element. Slower events, like load being higher than forecast, are not contingencies.

- Control generation (input locations and amounts) to avoid overloading transmission lines.
- Restart the system after it collapses because of failure to do one of the above.

ENERGY

The lack of energy storage and the varying needs for electricity result in volatile hourly energy prices. This is true in the vertically integrated environment where the marginal cost of power (system lambda) is optimized in economic dispatch. This also holds true in the restructured environment where hourly (and sub-hourly) markets are cleared based on energy bid prices.

The volatility of energy prices is important to the discussion of ancillary services. Opportunity costs in the energy markets drive ancillary service prices. Ancillary service prices are, consequently, typically more volatile than energy prices.

Energy is traded through long-term bilateral contracts and through hourly and sub-hourly markets in many regions. The NYISO operates a 5-minute energy market. Sub-hourly energy markets allow system operators to efficiently do a great deal of balancing of generation and load through energy markets without having to explicitly purchase additional control services. Sub-hourly energy markets may present an opportunity for responsive loads, which will be discussed in more detail later.

CAPACITY

Capacity is especially important in electric power systems because there is no energy storage. Sufficient capacity must always be available to serve immediate firm load requirements and to compensate for any system failures (contingencies) as shown in Figure 1. Vertically integrated utilities use long-term central planning to obtain needed capacity. In the restructured environment, the way capacity is obtained differs depending on the market structure.

A major debate continues concerning the advisability of creating explicit capacity markets versus having only energy markets. In California, where there are currently no explicit capacity markets, generators must make sufficient profit in the energy market to cover their capital costs. This means that the energy price a peaking unit receives must be very high if the unit runs only a few hours a year or it will go out of business. Some argue that this correctly values the cost of energy during the peak hours and that energy prices should not be capped. Unfortunately, capacity reserves, which may never be called upon to run, have no way to be paid. Market designers, especially in the northeast (and possibly FERC), worry that insufficient generation will be built to ensure reliability if generators must rely on only real-time energy markets. The alternative used by NYISO, PJM, and the Independent System Operator of New England (ISO-NE) is to run capacity markets that pay generators to be available and ensure that the marginal generators' fixed costs are covered. This reduces the volatility of the energy markets and ensures that capacity is available to serve expected load. Capacity payments can have the disadvantage of tending to draw low capital cost (and typically high operating cost) generators into the mix.

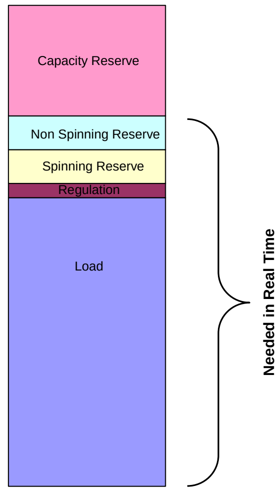

Figure 1. Capacity is required to meet load, provide regulation and contingency reserves, and to cover the unavailability of other generation and load forecast errors (capacity reserves).

Capacity markets can be problematic for loads. Generators can be available 24 hours a day, 365 days a year but most loads do not have uniform availability. A residential air conditioning load, for example, is not available continuously but it is inherently available when the power system is experiencing the peak summer load and capacity has value. Capacity market rules need to be carefully designed to allow responsive loads to participate appropriately. NYISO Installed Capacity/Special Case Resources rules will be discussed later.

ANCILLARY SERVICES

Energy and capacity are the basic products customers really use. But energy and capacity alone are not sufficient to reliably operate the power system. A series of ancillary services are required that provide the system operator with the resources needed to maintain the instantaneous and continuous balance between generation and load to manage transmission line flows. These services are required under normal conditions and when contingencies happen. Utilities have been performing these functions for a hundred years as part of their vertically integrated structure. Restructuring and the introduction of competitive generation markets has required that these services be clearly defined and monetized. Markets have been created for several ancillary services to minimize the cost of maintaining reliability.

Four of the five ancillary services (all except voltage control) deal with control of real power. The services are distinguished based upon response time, response duration, and response frequency. Faster, more frequent services get paid more (ancillary service prices will be discussed further in Section 4). Response duration, on the other hand, does not translate into higher service price. Response accuracy is not well quantified, but should be.[3] This will be discussed later as well.

The ancillary services are placed into three groups for discussion here: regulation that provides continuous response to balance generation and load under normal conditions, three services that provide reserves that stand ready to respond in the event of a power system contingency (10-min spinning reserve, 10-min non-synchronous reserve, and 30-min operating reserve), and voltage control or dynamic reactive power supply. Figure 2 shows how the ancillary services are differentiated in response time and duration, and Table 1 provides brief descriptions of all of the ancillary services.

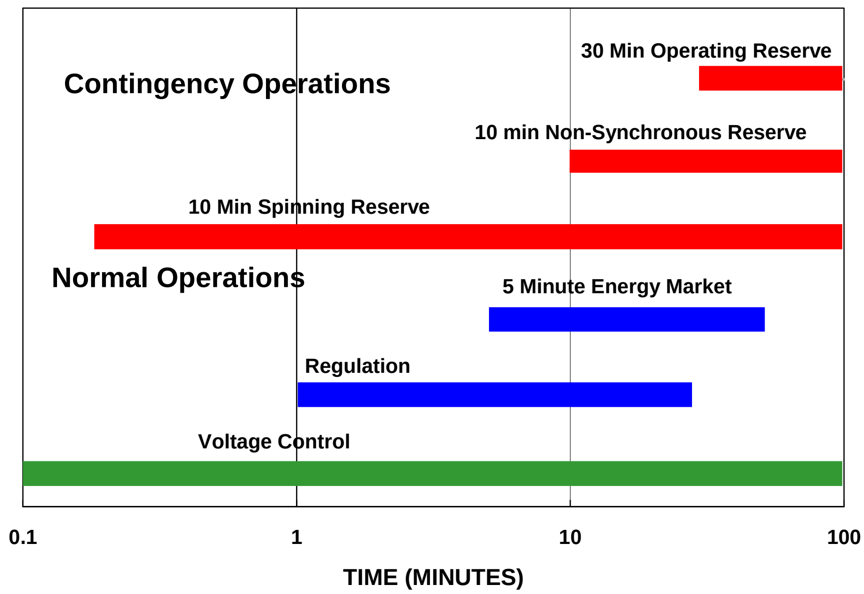

Figure 2. Response time and duration characterize required NYISO ancillary service response.

[3] Generators and loads with more accurate response should push for metrics that differentiate, and pay based on, the quality of response.

Table 1. Definitions of key ancillary services

Service	**Service Description**		
	Response Speed	*Duration*	*Cycle Time*
Normal Conditions			
Regulating Reserve	Online resources, on automatic generation control, that can respond rapidly to system-operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output to comply with control performance standards (CPSs) 1 and 2 of the North American Electric Reliability Council (NERC 2006)		
	~1 min	*Minutes*	*Minutes*
Load Following or Fast Energy Markets	Similar to regulation but slower. Bridges between the regulation service and the hourly energy markets. Supplied by the 5 minute energy market.		
	~10 minutes	*10 min to hours*	*10 min to hours*
Contingency Conditions			
10 Minute Spinning Reserve	Online generation or responsive load, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min to comply with NERC's Disturbance Control Standard (DCS)		
	Seconds to <10 min	*10 to 120 min*	*Hours to Days*
10 Min Non-Synchronous Reserve	Same as spinning reserve, but need not respond immediately; resources can be offline but still must be capable of reaching full output within the required 10 min		
	<10 min	*10 to 120 min*	*Hours to Days*
30 Min Operating Reserve	Same as non-synchronous reserve, but responds with 30 minutes; used to restore spinning and non-synchronous reserves to their pre-contingency status		
	<30 min	*2 hours*	*Hours to Days*
Other Services			
Voltage Control	The injection or absorption of reactive power to maintain transmission-system voltages within required ranges		
	Seconds	*Seconds*	*Continuous*
Black Start	Generation, in the correct location, that is able to start itself without support from the grid and which has sufficient real and reactive capability and control to be useful in energizing pieces of the transmission system and starting additional generators.		
	Minutes	*Hours*	*Months to Years*

Regulation

Regulation and load following, or fast energy markets, are the two services required to continuously balance generation and load under normal conditions. A typical daily load pattern with a morning ramp-up, double peak, and evening ramp down can be seen in Figure 3, which also shows the continuous, random minute-to-minute fluctuation in total system load that is superimposed on the predictable daily load. Regulation is the most expensive ancillary service and will be discussed in depth in this report.

Regulation is the use of on-line generation or responsive load that is equipped with automatic generation control (AGC) and that can change output quickly (MW/min) to track the moment-to-moment fluctuations in customer loads and correct for the unintended fluctuations in generation. Regulation helps maintain interconnection frequency, manage differences between actual and scheduled power flows between balancing areas, and match generation to load within the balancing area. Load following is the use of on-line generation, storage, or load equipment to track the intra- and inter-hour changes in customer loads. Regulation and load following characteristics are summarized in Table 2.

In New York, the PJM region, New England, and Ontario, regulation is a 5-min service, defined as five times the ramp rate in megawatts per minute. In Texas, it is a 15-min service, and in Alberta and California, it is a 10-min service.

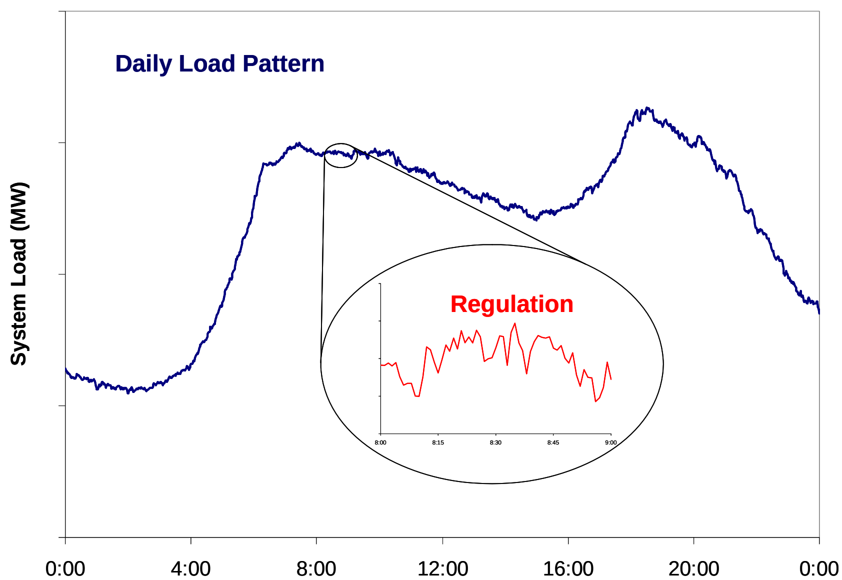

Figure 3. Regulation is a zero-energy service that compensates for minute-to-minute fluctuations in total system load and uncontrolled generation, while load following compensates for the slower, more predictable changes in load.

Balancing area operators have not needed to specifically procure load following; it is obtained from the short-term energy market with generators (typically) responding to real-time energy prices. Regulation, however, requires faster response than can be obtained from units responding to market signals alone. Instead, generators and responsive loads offer capacity that can be controlled by the system operator's AGC system to balance the power system. Typical balancing areas require 1–2% of the peak load in regulating capacity with larger balancing areas and balancing areas that operate sub-hourly energy markets requiring relatively less regulation than smaller balancing areas and balancing areas that only schedule generation and transactions hourly. The NYISO regulation requirement varies somewhat throughout the day and year (also on Sundays) as shown in Figure 4.

Table 2. Comparison of regulation and load following characteristics

	Regulation	Load following
Patterns	Random and uncorrelated	Highly correlated
Control	Requires AGC	Can be manual
Maximum swing	Small	10–20 times regulation
Ramp rate (MW/min)	5–10 times load following	Slow
Sign changes per unit time	20–50 times load following	Few

NYISO uses a reserve demand curve to limit prices in the event there is insufficient regulation supply. That is, if the NYISO needs 250 MW of regulation but there is only 200 MW offered, the price will not go to infinity. Instead the ISO limits the maximum price based on a “reserve demand curve.” The actual limit is \$300/MW-hr of regulation if NYISO is more than 25 MW short and \$250/MW-hr if the NYISO is less than 25 MW short. This curve is tied to an understanding of generator costs. The idea is that beyond a certain level, price is no longer a motivator: every generator that could supply would already be doing so. This is a good theory, but calculating responsive load opportunity costs is difficult, so the curve likely fails to motivate some demand response.

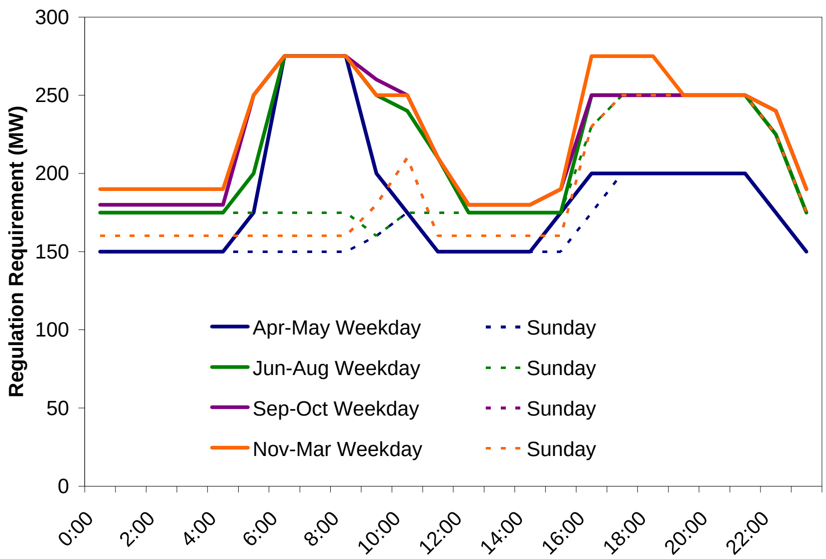

Figure 4. NYISO regulation requirement varies throughout the day and the year.

Balancing areas are not able and not required to perfectly match generation and load. The North American Electric Reliability Council (NERC) has established the control performance standard (CPS) to determine the amount of imbalance that is permissible for reliability purposes.[4] CPS1 measures the relationship between the balancing area's area control error (ACE) and the interconnection frequency on a 1-min average basis. CPS1 values can be either "good" or "bad." When frequency is above its reference value, under-generation benefits the interconnection by lowering frequency and leads to a good CPS1 value. Over-generation at such times, however, would further increase frequency and lead to a bad CPS1 value. CPS1, although recorded every minute, is evaluated and reported on an annual basis. NERC sets minimum CPS1 requirements that each balancing area must exceed each year.

CPS2, a monthly performance standard, sets control-area-specific limits on the maximum average ACE for every 10-min period. Balancing areas are permitted to exceed the CPS2 limit no more than 10% of the time. This 90% requirement means that a balancing area can have no more than 14.4 CPS2 violations per day, on average, during any month.

Energy Neutrality. Energy market impacts can be important for loads supplying regulation. Expected energy might be purchased in the day-ahead hourly market, for example, with differences between the expected average energy consumption and the actual energy consumption that results from following system operator commands settled at the sub-hourly energy market price. These two prices are typically close *on average*. But the average difference between the day-ahead and sub-hourly energy prices may not be relevant. One would expect the system operator to regulate the load down when the sub-hourly prices are high and regulate the load up when the prices are low. In the expected case, the sub-hourly energy price will average lower than the day-ahead energy price. This is not guaranteed, however. The system operator does not deploy regulation in response to energy prices but rather in response to balancing needs.

Similarly, regulation is expected to be a zero-energy service but this is not guaranteed either. That is, regulation-up should balance regulation-down over some reasonable period (a few hours). However, the system operator might regulate the load up or down for an extended period. This does not usually bother generators that sell regulation but it can be very detrimental to responsive loads and to storage resources. An aluminum plant that was regulated down for several hours, for example, could result in excessive cooling of the pot line and need to curtail regulation and increase load. Note that the NYISO has recognized this problem for storage-based resources and is beginning to address it. It may be necessary for load-based regulation resources to work with the NYISO to have their energy limitations accommodated as well.

[4] NERC is in the process of developing and testing three new metrics to replace the CPS and DCS metrics; control performance metric, disturbance control metric, and balancing authority area control error limit. These are not discussed here. The requirements for generators supplying ancillary services will not change significantly if the new metrics are adopted.

Response Accuracy. NYISO has a Regulation Performance Index that tracks how well a regulating resource follows instructions. Specifics on this index can be found in Section 6. Payments are reduced for poor performance. As shown in Figure 5, conventional generators, especially large thermal generators, do not follow control signals perfectly. The interconnected power system has been designed to accommodate this constraint, so reliability is maintained in spite of the deficiency. Balancing authorities typically require regulation capacity equal to 1–2% of the peak load to meet the CPS 1 and 2 limits, but the exact amount depends on the volatility of the load and the accuracy of the generators providing response; the more accurate the response, the less regulation required. System operators determine the amount of regulation required empirically. If the CPS scores are not high enough, operators increase the amount of regulation purchased. If CPS scores are too good, they decrease the amount.

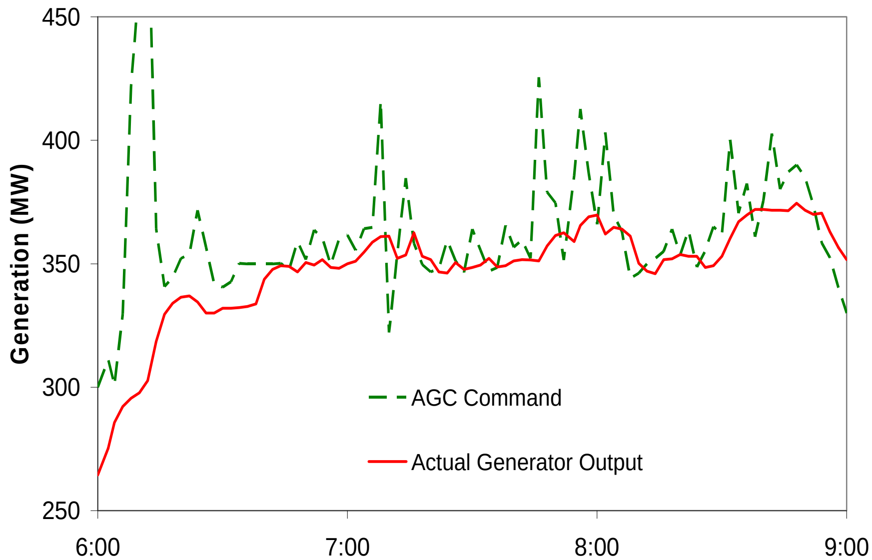

Figure 5. This coal fired power plant follows AGC regulation commands poorly.

Some regions do establish regulation certification criteria. NYISO and PJM, for example, use the response test shown in Figure 6 to certify generators to supply regulation. Generators are tested to ensure their ability to respond. Once certified, they can continue to supply (and be paid for) regulation unless the system operator notices that response has degraded unacceptably.

More accurate resources, such as solid-state based load control, would reduce the amount of required regulation. A 30,000 MW system that requires 250 MW of regulation capacity from its thermal generators might only require 200 MW of more accurate response capability to achieve the same CPS 1 and 2 scores. All things being equal, the 200 MW of more accurate response should be paid the same total dollars that the 250 MW of poorer response is currently being paid. The NYISO Regulation Performance Index may adequately reflect this difference and appropriately compensate accurate regulation providers.

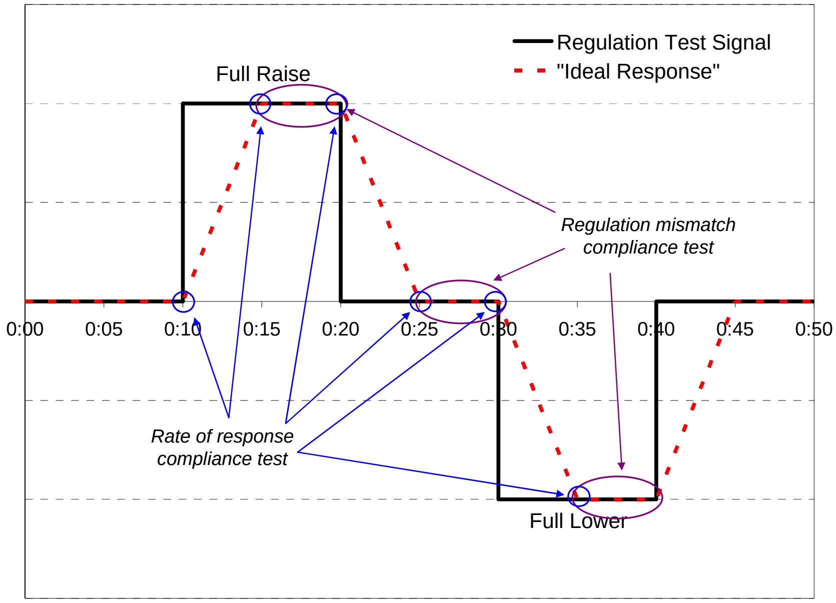

Figure 6. NYISO and PJM use a pass/fail test to certify and pay generators that supply regulation.

Services for Contingency Conditions

Generators and transmission lines can fail at any time. Contingency reserves restore the generation/load balance after the sudden unexpected loss of a major generator or transmission line. Power system frequency drops suddenly when generation trips, as shown in Figure 7. There is no time for markets to react. Frequency-sensitive generator governors and responsive loads respond immediately to stop the frequency drop. Ten-minute spinning and 10-minute non-synchronous reserves must restore the generation/load balance to return frequency to 60 Hz and ACE to zero within 15 min in order to meet NERC's disturbance control standard (DCS) requirements. Power systems typically keep contingency reserves available to compensate for the worst credible event (contingency). This is typically the loss of the largest generator or the largest importing transmission facility. In Texas, for example, the simultaneous loss of two nuclear plants is credible (as shown by the event recorded in Figure 7), so the Electric Reliability Council of Texas (ERCOT) requires over 2600 MW of contingency reserves.

The NYISO purchases three contingency reserves in a coordinated fashion: 10-min spinning reserve, 10–min total reserve (10-min spinning reserve + 10-min non-synchronous reserve), and 30-min total reserves (10-min total reserve + 30–min non-synchronous reserves). The NYISO also has locational requirements for the contingency reserves (regulation is a state-wide service) with a minimum amount of each reserve required on Long Island and in the Eastern region as shown in Table 3. The contingency reserves operate in a coordinated fashion to restore system balance as shown in Figure 8.

Table 3. NYISO contingency reserve requirements are coordinated in time and location.

Reserve Product	NYISO Total	Eastern NY	Long Island
10 Minute Spinning Reserve	600 MW (1/2 single largest contingency)	300 MW (1/4 single largest contingency)	60 MW(1/20 single largest contingency)
10 Minute Total Reserve	1200 MW (single largest contingency)	1200 MW	120 MW (1/10 single largest contingency)
30 Minute Reserve	1800 MW	1200 MW	270-540 MW

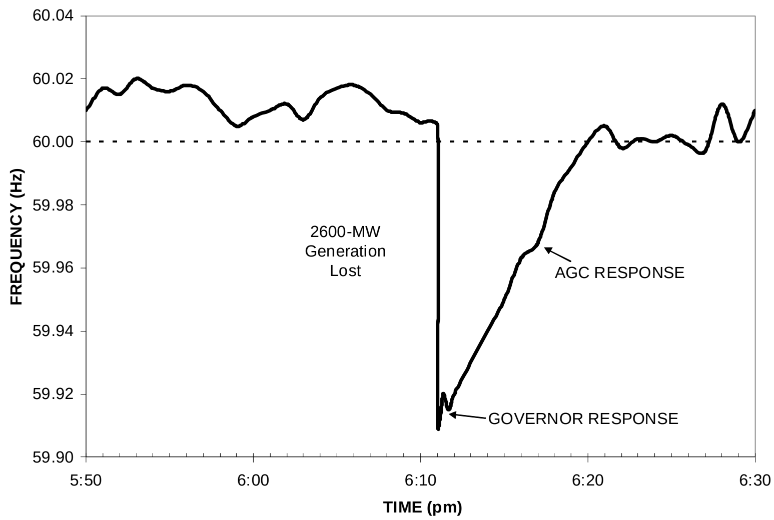

Figure 7. System frequency plummets in response to a major loss of generation and the generation/load balance must be restored quickly.

10-Minute Spinning Reserve. Spinning reserve is supplied by generation and responsive load that is on-line, less than fully loaded, begins responding immediately, and is fully responsive within ten minutes. The resource continuously stands ready to respond to a major loss of generation or transmission and can be deployed autonomously if system frequency falls or in response to a system operator's command. The resource must have a governor to sense and respond to frequency drops and telecommunications (typically AGC) to respond to system operator deployment commands. In NYISO, the AGC calculates an ACE. This error is then apportioned to the selected Regulation Resources in the ratio of the amount of their planned Regulation Service. Spinning reserves must be capable of sustaining the response, typically for two hours, though system operators try to relieve spinning reserves much sooner in order to be ready for the next contingency. The average contingency reserve deployment event lasts 11 minutes in New York.

A generator can be limited in ramp rate (MW/min) or in total available generation. Metrics are not well defined; there are no specifics about how much of the generation must respond "immediately" or even what "immediately" means. Similar to regulation, spinning reserve is a service where responsive load has a technical advantage in terms of speed and accuracy of response. ISO market rules do not yet recognize or pay for the superior load response, but they might if a strong case were made.

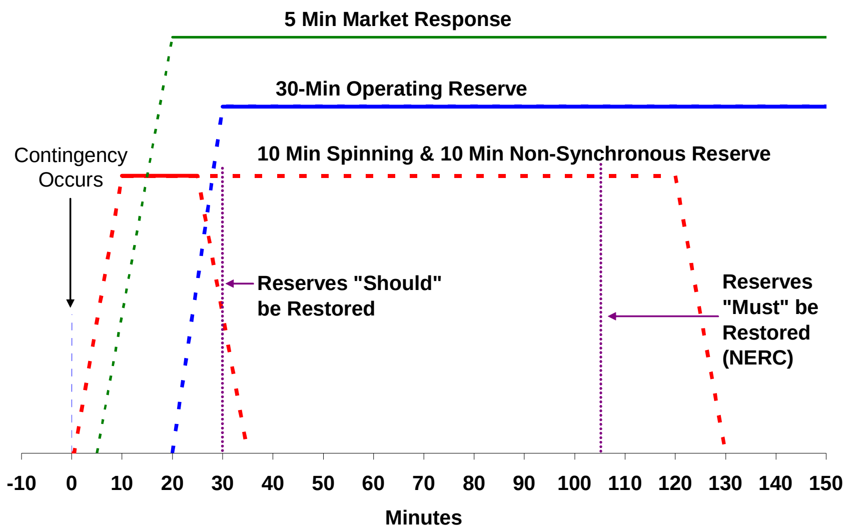

Figure 8. A series of coordinated contingency reserves restore the system generation/load balance immediately following a major contingency.

10-Minute Non-Synchronous Reserve. Non-synchronous reserve is similar to spinning reserve in that the resource must be fully responsive within 10 minutes and must be capable of sustaining the response for 2 hours. The resource does not have to be on-line and spinning, it does not have to begin responding immediately, and it does not have to be frequency responsive. The resource must have telecommunications to respond to system operator deployment commands. Any load that can provide 10-minute non-synchronous reserve should strive to supply 10-minute spinning reserve instead. The response requirements for loads are nearly identical, and the spinning reserve price is significantly higher.

30-Minute Operating Reserve. Thirty-minute operating reserve is similar to 10-minute non-synchronous reserve with a slower required response. This reserve must be fully deployed in 30 minutes and must be capable of sustaining that response for 2 hours. On line generation, off line generation, and responsive load can provide this reserve. The reserve responds to system operator commands to deploy and to restore and must have telecommunications, but AGC is not necessary.

Contingency Reserve Deployment Frequency and Duration. While the power system must have sufficient contingency reserves constantly standing by ready to immediately respond to the sudden failure of the largest generator or transmission facility, actual deployment frequency depends on the actual failure rates. Systems also differ in how often they call on reserves or on what types of events justify reserve deployment. NYISO, for example, uses contingency reserves relatively frequently; 239 times in 2002.[5] ISO-NE and CAISO use contingency reserves much more sparingly; only 19 and 26 times respectively in 2005. The frequency of reserve deployment depends on the entire generation mix, the number of large units, and their failure rates. It also depends on the mix of responsive resources in the economic pool that are also available for system operator dispatch in the event of a smaller contingency. In all three cases, the response *duration* was typically short (~10 min) as shown in Figure 9. But it is critical that the reserves be *capable* of longer response in the event of a truly serious disturbance.

Frequency Responsive Reserve. It has been recognized for some time that there is little technical justification for the specifics in the spinning reserve requirements. This is most apparent in the western U.S. where the Western Electricity Coordinating Council (WECC) requires maintaining contingency reserves equal to 7% of the load served by thermal generation and 5% of the load served by hydro with at least half being spinning. There is no technical basis for the 5%, the 7%, or the half. The WECC Reserve Issues Task Force is attempting to address this lack by proposing a new frequency responsive reserve (FRR), which does have technical justification. The proposed criteria are currently undergoing the WECC approval process.

FRR would replace, and be very similar to, the current spinning reserve. It is designed to address generator response in the 30 second time frame. Future work will address 10–20 second response. All of WECC would be required to carry only 3200 MW of FRR, an amount equal to the largest credible (category C) double generator contingency that WECC plans to be able to withstand without under-frequency load shedding (59.5 Hz). The total requirement would be allocated to balancing authorities based on a peak load ratio share. Each balancing authority would still be required to carry enough contingency reserves (FRR plus non-synchronous) to cover the balancing authority's largest single contingency.

FRR has historically been a part of spinning reserve and the required additional governor response of all on-line generators. NERC is considering specifically calling out this reserve as a named (and presumably paid for) ancillary service. This could be a profitable service for some responsive loads since the full load response is often available essentially instantaneously. Generators, on the other hand, take a full 10 minutes to ramp up spinning reserve so only a fraction of their spinning reserve response is available as FRR.

[5] NYISO use of contingency reserves was reverified with 2008 data with essentially the same results.

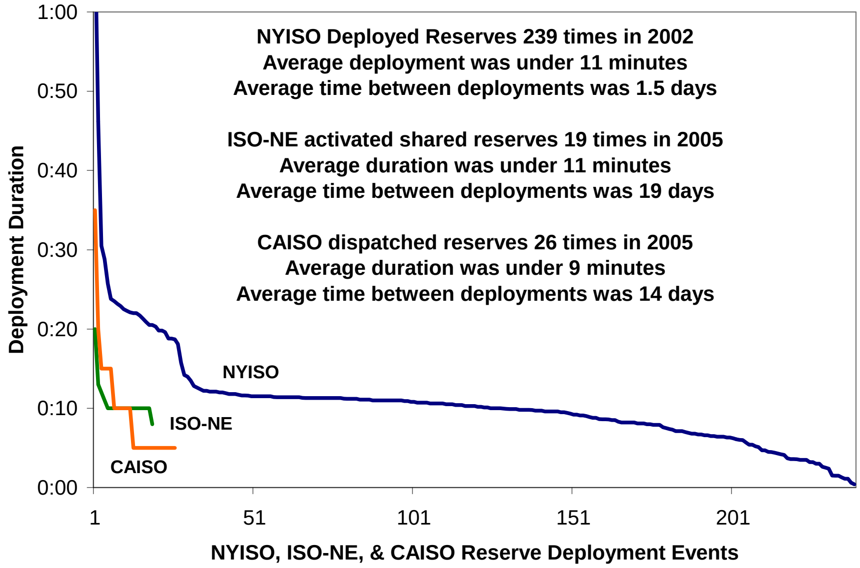

Figure 9. System operators differ in how often they deploy contingency reserves but events are usually relatively short.

VOLTAGE CONTROL

Unlike the other ancillary services listed in Table 1, voltage control is not a real-power service. Instead, it involves the control of reactive power to maintain acceptable voltages throughout the power system under normal and contingency conditions. Reactive power is measured in VARs (volt amps reactive) or MVARs (millions of volt amps reactive). The units are similar to watts and MWs except that the voltage and current are 90^o out of phase. Power system voltage is sensitive to, and controlled by, the injection and withdrawal of reactive power. Voltages must be maintained within a fairly tight range throughout the power system to protect customer and utility equipment and to prevent voltage collapse. Voltage that is too high can destroy equipment by breaking down insulation. Voltage that is too low can make motors stall and equipment overheat. Voltage collapse can occur when a cascading drop in voltage suddenly spreads throughout a region. To protect against these failures and to compensate for the reactive power that loads and the transmission system itself consume, the system operator must have reactive power resources available.[6]

Various pieces of equipment on the transmission and distribution system provide relatively inexpensive voltage control and reactive power. Capacitors, inductors, and transformer tap changes are all used for voltage regulation. But these transmission based solutions are slow to respond. The worst thing about using capacitors is that the reactive support they provide drops with the square of the voltage, which means they provide less support when they are needed most.

[6] Greater detail can be found in: B. Kirby and E. Hirst 1997, *Ancillary-Service Details: Voltage Control*, ORNL/CON-453, Oak Ridge National Laboratory, Oak Ridge TN, December available at www.consultkirby.com.

The power system requires a significant amount of dynamic reactive support to quickly control voltage especially during contingencies. Synchronous generators are excellent suppliers of dynamic reactive power. They inherently produce more support during faults and they can be autonomously controlled to maintain local voltage to a coordinated schedule set by the system operator. Reactive support from generators is so important for reliability that the FERC Order 2003 (and supplemental Orders 2003-A, 2003-B, and 2003-C) covering Large Generator Interconnection Procedures (LGIPs) and Large Generator Interconnection Agreements (LGIAs) requires that synchronous generators be designed and operated such that they can supply reactive power within the range of 0.95 leading to 0.95 lagging while maintaining full rated real power output, “unless the Transmission Provider has established different requirements applicable to all Interconnection Customers.” (FERC Order 2003, Article 9.6.1) The NYISO provides an administrative compensation of $39.10/KVar per year to all generators that demonstrate reactive supply capability through annual performance testing.

A synchronous generator’s ability to produce reactive power is related to its ability to produce real power but not linearly, as shown in Figure 10. A similar interaction between real and reactive capabilities exists for solid-state-based devices such as variable frequency motor drives and solid-state power supplies. Reactive power costs and compensation will be discussed in Section 4, but it is important to point out here that the electrical generator or solid-state load must be designed to be larger (MVA) than the real power (MW) requirements would dictate, specifically to provide reactive support capability while operating at full real power. The vertical dotted constant-power line in Figure 10shows that this machine is capable of operating at a 0.85 power factor when the prime mover is producing full real power or the load is meeting full capacity.[7] If more reactive power is required from this resource, it is necessary to reduce the real power consumption/production. The decision concerning how much extra reactive power capability is built into the resource is made at design time and impacts the resources’ usefulness to supply dynamic reactive support for the rest of its life. It is basically a question of initial capital cost and expected long-term compensation.

The need for dynamic reactive power changes from time to time and from location to location. Changes in transmission-system loading change the reactive power consumption of the transmission system itself. As loads’ real power requirements change, the reactive power requirements change as well. Reactive power requirements are location specific because the inductive impedance of the transmission system is much greater than the resistance —VARs don’t travel well.

[7] Power Factor is the ratio of real power (MW or watts) to apparent power (MVA or volts times amps). It can be expressed in per unit (0.85) or percent (85%).

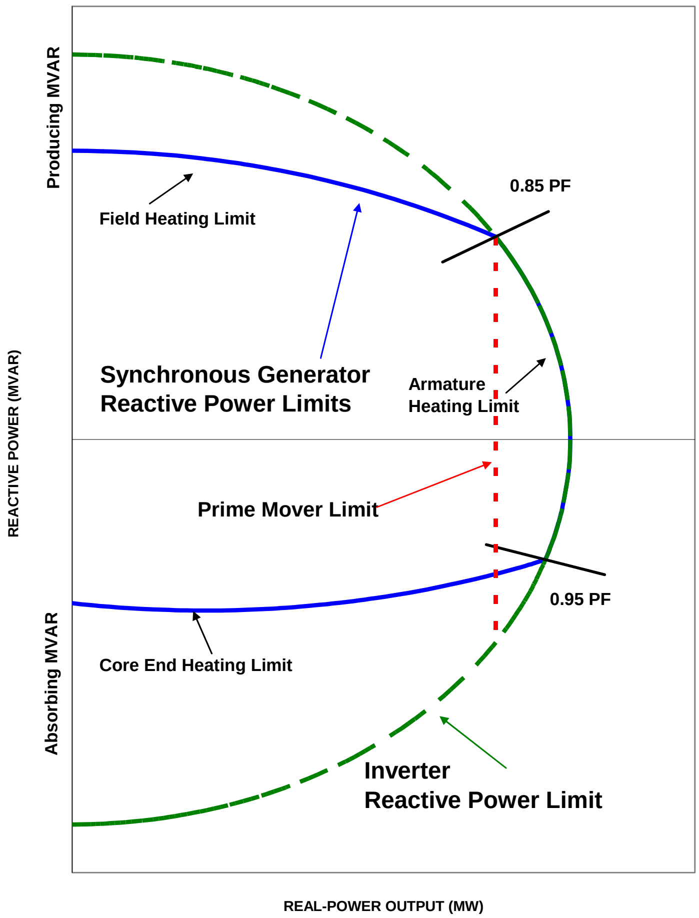

Figure 10. Real and reactive power production capabilities are interrelated in both synchronous generators and power electronics devices.

BLACK START

Unlike many other engineered systems, the four North American power system interconnections are designed to be in continuous *synchronized* operation. Individual pieces of equipment are taken out of service for maintenance, but each interconnection as a whole is designed to run without interruption.[8] Nevertheless, the power system must be prepared for the rare occasions when all or a major portion of the system is forced out of service. This might be the result of a particularly severe disturbance resulting in the loss of stability and the need for many generators to shut down. If this occurs, the system must be able to be restored to normal operations as quickly as possible.

One critical resource that must be available when restoring a power system to service is black start generators: generating units (often hydro and combustion turbine units) that can start themselves without an

[8] B. Kirby and E. Hirst 1999, *Maintaining System Black Start in Competitive Bulk-Power Markets*, American Power Conference, Chicago, Illinois, April.

external electricity source and can then energize transmission lines, restart other generating units, and ultimately restore service to customers.

Selecting the black start generators is somewhat location dependent. Black start generators must be electrically close enough to the other units they are to help restart to be able to energize the transmission lines connecting the two plants and control voltages at both ends. The black start units must also have sufficient capacity and ramping capability to be able to provide the restart power required by the other units. The system operator determines how many units within the control area must have black start capability, where they are to be located, and how to use them in the event of a blackout.

Note that this is one ancillary service that requires the *production* of real and reactive power. Curtailing load is not sufficient. In fact, by definition, all load is curtailed at the time when black start is required. Loads are not physically capable of providing black start unless the load has a behind-the-meter generator. Even in that case, the generator must be fairly large (at least 50 MW) and have both real and reactive power control. If a load has a suitable generator in the correct location it can negotiate supply of black start capability with the NYISO.

3. LOAD RESPONSE CAPABILITIES THAT ARE VALUABLE FOR POWER SYSTEM RELIABILITY

Loads can interact with the power system in range of ways with a range of economic consequences. The simplest way for loads to behave is to ignore the power system and consume energy whenever required by whatever function it is trying to accomplish (rolling steel, air conditioning a house, moving a subway train, etc.). For many loads, this will be the least-cost method to conduct their business. The cost of changing energy consumption may outweigh any benefit of reduced energy price. Operating subway trains only at night, for example, would reduce the cost of electricity but would also excessively inconvenience commuters. Loads that are able to change energy consumption have a range of options, all with different costs to the load and different benefits for the power system. For loads that are able to adjust consumption, there are five basic ways to provide useful response as shown in Figure 11. All of them have some impact on power system reliability; some have a greater impact than others.[9]

- **Energy Efficiency** programs reduce electricity consumption and usually reduce peak demand
- **Price Response** programs move consumption from day to night (real time pricing or time of use)
- **Peak Shaving** programs require more response during peak hours and focus on reducing peaks every high-load day
- **Reliability Response** (contingency response) requires the fastest, shortest duration response. Response is only required during power system "events" – *this is new and slowly developing*
- **Regulation Response** continuously follows the power system's minute-to-minute commands to balance the aggregate system – *this is very new and may have the potential to dramatically change production costs, especially for aluminum and chlor-alkali*

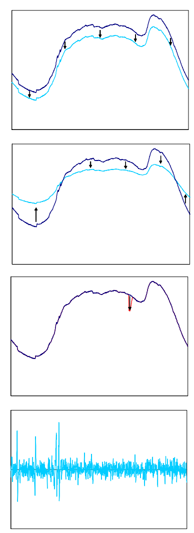

Figure 11. All five basic types of demand response impact power system reliability.

Energy efficiency reduces consumption during all hours and typically reduces the need for generation and transmission. It is not focused on times of greatest power system stress and compared with more directed alternatives, may not provide as cost effective a reliability response to specific reliability problems. Price

[9] B. Kirby, 2006, *Demand Response For Power System Reliability: FAQ*, ORNL/TM 2006/565, Oak Ridge National Laboratory, December

responsive load and peak shaving both target specific hours when response is desired: the former facilitates voluntary market response to price signals while the latter utilizes direct control commands. Both types can be used to address capacity inadequacy caused by a lack of generation or a lack of transmission. Reliability response (contingency response) and regulation specifically target power system reliability needs through the supply of ancillary services and offer the greatest reliability benefit per megawatt of load from loads that are capable of providing these types of response.

As seen in Section 2, the ancillary services that provide power system reliability are based on maintaining the generation/load balance. This balance must be maintained instantaneously and continuously. Few resources (generators or loads) can economically provide the full range of required response so that requirements are broken down into separate ancillary services, primarily distinguished by the required response speed, duration, and frequency. Traditionally, power system operators have controlled generation to maintain the balance, but technically the balance can be maintained by adjusting either generation or load.[10] Advances in communications and control now make it possible for loads to provide response that is equivalent to or better than that provided by generators.

The ability of a load to provide response services to the electric power system is constrained by the energy needs of the load, the speed of the communications and controls, the response speed of the load itself, and the load's "energy storage" capability that constrains response duration.

LOAD ENERGY REQUIREMENTS

The vast majority of power system reliability response needs involves increasing generation or decreasing load. The sudden failure of a generator must be met with a sudden increase in output from other generators or a sudden decrease in net load. A poor weather forecast that underestimates summer afternoon temperature will result in underestimating peak load and the possibility that insufficient generation will be available to meet the load requirement. Additional generation is required or load must be reduced. Overestimating summer temperature and over-scheduling generation may not be economical but is generally not a reliability problem.[11] A load must be consuming energy to be able to provide useful reliability response. Further, the load must have the flexibility to curtail its energy consumption.

Loads with energy requirements that are coincident with the overall power system (residential air conditioning, for example) are disadvantaged in that they require energy exactly when it is most expensive.

[10] Under-frequency load shedding has always been used to maintain reliability, but this is an uncompensated and involuntary use of load response and is only deployed as a last resort to prevent system collapse.
[11] Excess generation at night can be a reliability problem (minimum load) but this is usually dealt with through energy prices rather than through reserve deployment.

This also means that the load is consuming energy when the contingency reserves are most expensive, so the load is potentially available as an ancillary service provider. Industrial loads that can minimize consumption during high price times may reduce their ability to sell contingency reserves when those services are most expensive.

COMMUNICATIONS AND CONTROL SPEED

Communications and control are the critical characteristics that determine if a load is capable of providing reliability reserves. Load must be controllable if it is to supply reliability services to the power system. The control must be fast and accurate. The load must also have a way to receive deployment commands from the power system operator. The required response speed and duration depend on which reliability service is being provided.

Communications and control for peak reduction are not required to be fast. Loads are provided with hours of warning. E-mail and phone calls, as well as manual control, are adequate. Automating response is an economic question, not a reliability/speed requirement.

Contingency reserves and regulation require fast communications. Communications with generators providing contingency reserves and regulation occurs every two to six seconds with automatic response initiated immediately. Conventional power system supervisory control and data acquisition (SCADA) and AGC are likely adequate for large industrial load control. While the cost is relatively high per installation, it is spread over a large response. Conventional SCADA methods are likely not cost effective for large aggregations of small loads, and lower cost communications are required.

LOAD RESPONSE SPEED

Response speed is an area where load response typically greatly exceeds generation response capability. The most important and highest cost ancillary services require the fastest response. Regulation is a minute-to-minute service; spinning reserve requires response to begin immediately with full response in ten minutes. Many loads provide response by simply shutting down. A circuit breaker is opened providing full immediate response. Generators have to ramp up response and cannot match this reliability benefit.

Some loads do require advance notice in order to provide response. Work has to be rescheduled; industrial processes have to be adjusted, and additional manpower may be required. These loads can provide day-ahead response and peak load reduction. They cannot provide contingency reserves or regulation.

RESPONSE DURATION AND ENERGY “STORAGE”

Response duration is limited by the storage capability of the load. Storage could be in terms of electricity itself but it rarely is. An uninterruptible power supply (UPS) protects a computer load, for example, by buffering the electricity supply. Batteries within the UPS store enough electricity to allow the computer to keep running for many minutes even if the utility supply is interrupted. Conceptually, a UPS could be used to allow a load to sell ancillary services to the power system while continuing to operate normally. The size of the battery would determine how long the load could sustain the curtailment response.

While storage is critical in determining how long response can be maintained, it is typically not provided on the electric-power-input side of the process. Storage is typically provided in the form of the product the process is producing or in an intermediate product within the process. Commercial and residential air conditioning and refrigeration can provide excellent fast and accurate spinning reserve response, for example, by using the thermal mass of the conditioned space for storage. Ten to 30 minute response duration can easily be provided with complete curtailment of the compressor load. The inherent storage is typically not sufficient to provide multi-hour peak reduction, however. Peak reduction can be provided but it is necessary the air conditioning must be cycled on and off, which provides somewhat reduced multi-hour response.

Industrial processes can often provide instantaneous or fast response. The storage medium may be final product, intermediate product, or an ancillary characteristic of the internal process. In aluminum smelting, for example, the aluminum product is typically not the limiting element; that can fairly easily be stored. Instead, it is the thermal balance of the aluminum smelting pot line that is critical. The pots must remain hot enough to keep the cryolite molten. Pot heat comes from the electricity that is producing the aluminum, so interruption times are limited to a few hours at most. Control can be very fast and accurate. Aluminum production is typically base loaded so the resource is always available. Only the response duration is limited. Each aluminum smelter has its own operating characteristics that will determine the maximum duration of a peak reduction, but it is an ideal supplier of spinning reserve, based on the energy storage available within the process.

Each industrial process has to be examined to determine the capabilities and limitations. Control may be of secondary variables as well. Uranium enrichment (admittedly a unique process but the concept applies more generally) is controlled by adjusting the amount of gas in the process, not by directly controlling the energy consuming compressors. Turning compressors on and off creates expensive maintenance problems. Adjusting the gas inventory is fast, relatively easy, and cheap. In most processes it pays to examine secondary effects.

Many pumping loads are good candidates (water, natural gas, and other gases) for supplying ancillary services. Any industrial process with some manufacturing flexibility is a good candidate (cement, paper, steel, aluminum, refining, air liquefaction, etc.). The list is endless.

Regulation is more difficult for loads to supply, but electrolysis loads such as aluminum and chlor-alkali production appear to be excellent candidates.

ASSESSING RESPONSE VALUE

Several factors determine how much value a load can obtain by offering response to the power system. One factor is the price the power system pays for actual response, typically in $/MWH. Another is the amount the power system pays for the load to stand ready to respond, typically in $/MW-hr. A third factor is the cost the load incurs to either respond or to stand ready. Time series modeling of the power system and the load can determine which services should be offered each hour. Time series modeling is beneficial because energy and ancillary service prices are volatile, as are load conditions. Fortunately, hourly prices for energy and the ancillary services are available going back several years. Typically, the mix of energy procurement and the specific response service being sold varies from hour to hour and season to season.

4. ANCILLARY SERVICE AND ENERGY PRICES

The NYISO (and most ISOs) conduct hourly markets for four ancillary services: regulation, 10-minute spinning reserve, 10-minute non-synchronous reserve, and 30-minute operating reserve.[12] Energy markets clear in three time frames as well: day-ahead-hourly, hour-ahead-hourly, and every 5 minutes.[13] Market prices are posted for each market interval and are available to create multi-year time-series data sets for analysis.

Regulation is the most expensive ancillary service followed by 10-minute spinning reserve, 10-minute non-synchronous reserve, and 30-minute operating reserve.[14] Response speed is valued more highly than response duration—this is important, and fortunate, for responsive load.

Cost drivers for each ancillary service will be discussed below in a bit more detail but all are driven primarily by opportunity cost. In order to sell into the ancillary service markets, generators must withhold capacity from the energy market. The cost the generator has to charge (or bid) to supply a reserve service is based primarily on the difference between the generator's production cost and the energy sale price for that hour. A generator with a production cost of $50/MWH, for example, would bid $10/MW-hr to sell spinning reserve if the energy price was $60/MWH. At any price higher than $10/MW-hr for spinning reserve, the generator makes more profit by forgoing the energy sale and selling spinning reserve. Conversely, at any price below $10/MW-hr for spinning reserve, the generator would lose money by staying out of the energy market.

One consequence of this linkage between energy and ancillary service markets is that ancillary service prices are inherently more volatile than energy prices. Contingency reserve prices, for example, are typically zero at night when numerous generators are at minimum load and have capacity available at essentially no cost.

Another important consequence for responsive loads selling ancillary services is that ancillary service prices may collapse if loads, with near zero opportunity cost, begin setting ancillary service prices. Responsive loads typically do not want to take 100% of the ancillary service market. They want a generator, with opportunity costs, to continue to set the market clearing price.

---

[12] Reserve names and definitions vary slightly from ISO to ISO. Ten-minute non-synchronous reserve is often called non-spinning reserve. Some ISOs do not have the 30-minute operating reserve and others have a 60-minute operating reserve. The slower reserve is sometimes called "supplemental" operating reserve or "replacement" reserve. Still, the requirements are essentially the same.

[13] Most ISOs operate sub-hourly markets, or are instituting them. ERCOT operates a 15 minute market but is working on replacing it with a 5 minute market.

[14] Some markets split regulation into regulation up and regulation down. This distinction is largely semantic rather than technical. Prices in split markets can be compared with combined markets simply by adding the up and down prices.

Note that the price unit for reserves is $/MW-hr. This is because the resource is selling 1 MW of *capacity* (not energy) for 1 hour. The resource is standing ready to produce or curtail but is not necessarily producing or curtailing. In fact, if the resource does deliver any energy during the hour, the cost of the energy will be settled separately, either at the resource's cost or at the spot energy price. Typically the energy component of the ancillary services is not major. This terminology is not universal but it does clearly distinguish the energy and capacity components.

REGULATION COST DRIVERS

The direct cost for a generator supplying regulation includes a degraded heat rate and increased wear and tear on the unit. The dominant expense, however, is the lost opportunity cost associated with maneuvering the generator in the energy market so that it has capacity available to sell in the regulation market. For example, a 600-MW generator with a full power energy production cost of $15/MWh would have to bid $27/MW-hr of regulation if the energy market were clearing at $30/MWh. This is to compensate the generator for the lost profit in the energy market when it reduces output in order to create maneuvering room to supply regulation and to compensate for the reduced efficiency (increased heat rate) associated with the remaining output's still being sold into the energy market. Figure 12 shows how a generator's cost (and bid price) to supply regulation depends upon the current energy price. Note too that this generator is limited to supplying only about 12 MW of regulation (~2% of its rated capacity). This is because regulation is a quick service and the unit ramp rate, rather than the total available capacity, limits the peak amount of regulation it can provide. For this reason, regulation is generally spread across several generators. Opportunity costs similarly dominate contingency reserve prices.

There is also an opportunity cost when the energy market price is below the generator's marginal production cost. When energy prices are low (typically at night) and generators are at minimum load, they incur a cost for running above minimum load in order to supply down regulation. For example, a generator with a 150-MW minimum load and an energy production cost of $18/MWh would have to bid $64/MW-hr of regulation if the energy market were clearing at $14/MWh because it would be losing $4 for each of the 162 MWh it must sell into the energy market to get its base operating point high enough to provide room to regulate down.

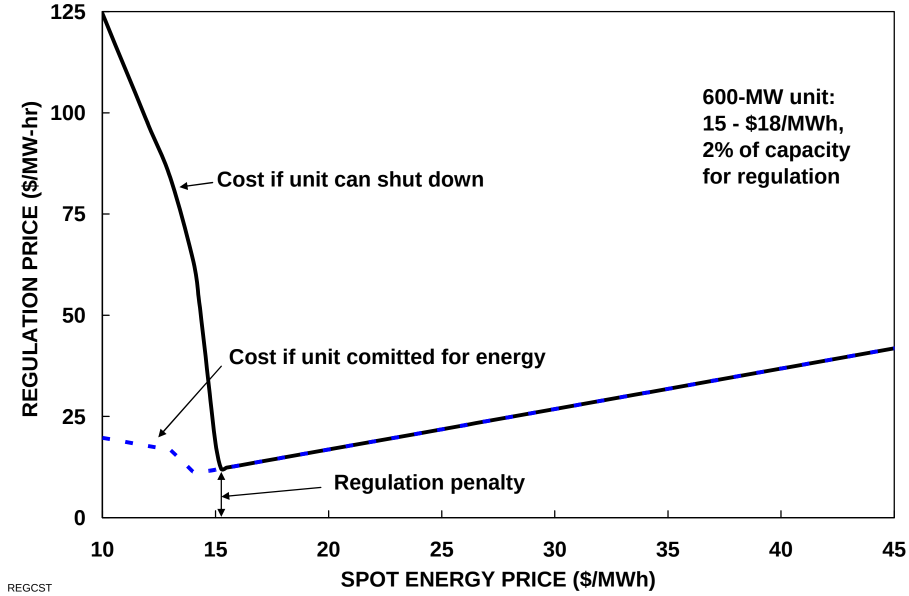

Figure 12. Regulation costs are dominated by generator opportunity costs. Cost at night can be higher than during the day.

Note that a load providing regulation does not incur exactly the same costs. The load likely does incur an efficiency penalty. It also incurs a higher capital cost because it must install equipment rated for the higher peak power to still consume same average energy. It will also likely incur opportunity cost but those opportunity costs will be related to shifting production of whatever it is that the load normally produces or the service the load provides. The load must calculate the efficiency penalty and the opportunity cost penalty itself and include them in the ancillary service bid price. The capital cost associated with the higher capacity equipment is not included in the ancillary service bid (it is a sunk cost by bid time).

Calculating load opportunity costs is much more difficult than calculating generator opportunity costs. This is especially true for the ISO since the ISO likely has little expertise concerning the load's area of business. If loads become dominant providers of ancillary services, the ISO will need to find ways to ensure that prices are appropriate and fair.

CONTINGENCY RESERVE COST DRIVERS

Contingency reserve cost drivers are essentially a subset of the regulation cost drivers. Because contingency reserves deploy infrequently, there is typically no significant degradation in generator or load efficiency and no increased wear-and-tear on the resource. Only the opportunity costs are incurred because the resource must withhold capacity from the energy market.

HOURLY AND SUB-HOURLY ENERGY MARKETS

The NYISO operates a series of energy markets that facilitate matching generation and load. While most energy is arranged through bilateral contracts, public markets provide price transparence for all market participants. A series of energy markets with successive market closing times facilitates balancing loads' needs with generation availability. Successive markets allow participants to refine their market positions. The NYISO conducts day-ahead hourly markets, hour-ahead hourly markets, and 5-minute real-time markets. Figure 14 presents the market-clearing time line for the NYISO and compares it with ISO-NE.

Locational Prices

Energy and ancillary service prices within ISOs vary from location to location. Transmission congestion can create significant price differences when there is insufficient transmission capacity to serve all of the load with the lowest cost generation. Transmission losses create price differences from bus to bus even when there is no transmission congestion.

The total amount of data can be overwhelming. NYISO provides prices for 11 internal zones (Figure 14), 4 external zones and 468 generator locations. Removing losses and congestion leaves the price for the system's "reference bus". The reference bus price is no more representative than any other location but it is a convenient choice when there is no compelling reason to select any other location.

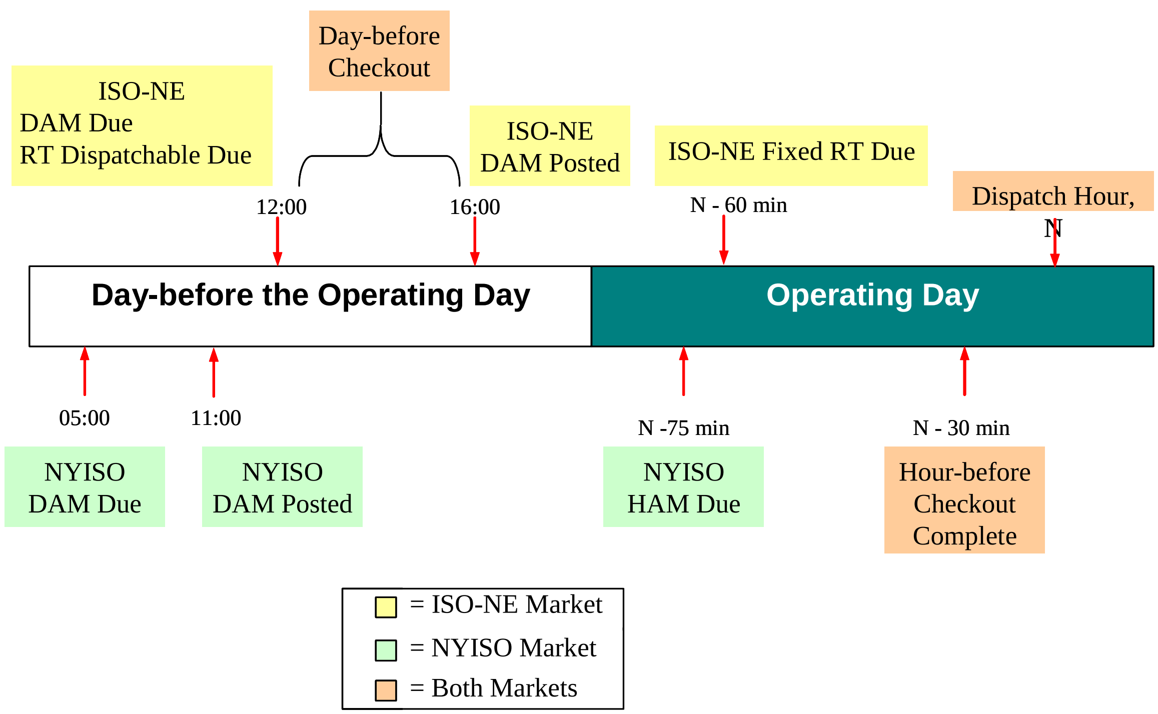

Figure 13. Market clearing times for ISO-NE and the NYISO.

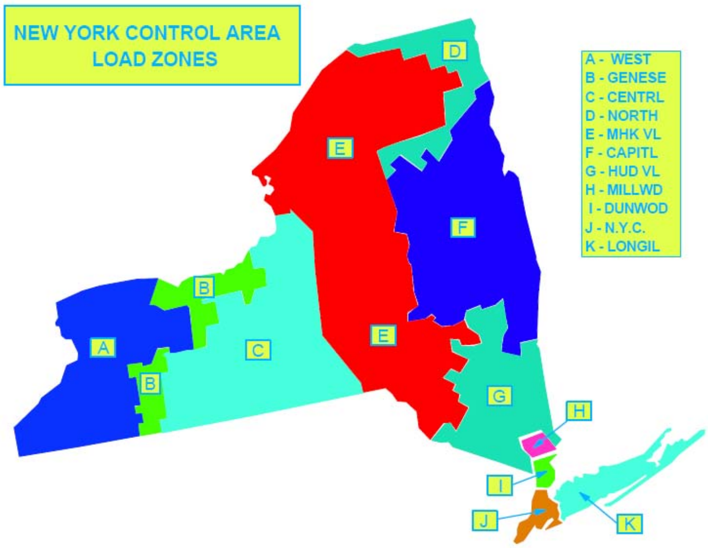

Figure 14. NYISO posts prices for 11 internal zones.

When using ancillary service and energy price data in analysis it is important to keep the location variability in mind and to always test to be sure that the results are not ignoring important congestion considerations.

CO-OPTIMIZATION

Co-optimization, while excellent for generators is typically problematic for responsive loads and storage devices providing ancillary services. Co-optimization is a technique used by the best ISO market operators (including NYISO) to determine the most profitable mix of energy and ancillary services each *generator* should sell and, simultaneously, what is the lowest cost mix of services the power system should purchase from each generator. It relies on the assumption that a generator that can supply regulation can also supply spinning reserve, non-synchronous reserve, supplemental operating reserve, and energy. Speed of response, not response duration, is assumed to be the limiting capability.

First let us consider why co-optimization works so well for generators. Provision of energy and the three reserve services are interrelated. It can be difficult to determine how much capacity to offer into each market and at what price. A generator will naturally want to maximize its profits and sell as much capacity as it can into whichever reserve market is paying the highest price. Alternatively, it will want to forgo the reserve markets and sell as much as possible into the energy market if that is providing higher profits. The earliest market designs cleared the hourly energy and reserve markets in sequence. The system operator first selected the least expensive set of generators to supply energy. Generators to supply regulation were selected next, followed by spinning reserve, non-synchronous reserve, and supplemental reserve. The thinking was that the energy market went first because it was the highest volume and the most important economically. Among the reserves, it was reasoned, it was necessary to procure the most technically demanding services first, followed by services that more generators can provide. This was done to prevent the 30-minute operating reserve market, for example, selecting generation that was needed to supply regulation. This makes sense because a generator that can supply regulation can probably supply spinning reserve, non-synchronous, and 30-minute operating reserve but the reverse is not necessarily true.[15] Unfortunately this can lead to perverse results in some cases. There may be sufficient low cost generation to supply the spinning reserve need, for example, but not enough to meet the need for 30-minute operating reserves once that market is finally cleared. In that case, the 30-minute operating reserve market will clear at a higher price than the spinning reserve market. This result is undesirable because it provides an incentive for the technically more agile generators to withhold capacity from the spinning reserve market and offer it into the 30-minute operating reserve market.

A solution was developed whereby the generators simply offer their capabilities and their costs. The system operator then co-optimizes the energy and ancillary service markets, guaranteeing each generator the maximum profit and the system the lowest combined cost. California's initial move in this direction was called the "Rational Buyer" which allowed the system operator to substitute "higher quality" reserves for "lower quality ones" but to pay the higher quality price. California is now going to a full co-optimization. NYISO, ISO-NE, PJM, ERCOT, and CAISO all perform co-optimization of some form.

Unfortunately co-optimization is ruinous to loads and storage devices trying to provide ancillary services. These resources can be ideal ancillary service providers, being both fast and accurate. They are often energy limited, however, and are unable to provide hours of continuous energy response. A load that risks being co-optimized into the multi-hour energy market simply will not offer to sell spinning reserve or regulation. This denies the load as a significant revenue stream and denies the power system as an excellent reliability resource.

Technically there is a fairly simple solution: allow resources to declare themselves unavailable for co-optimization. No generator capable of supplying energy as well as ancillary services will do so since it can only reduce the generator's profit. So co-optimization will continue to be used for most market participants. Some system operators have refused to allow resources to opt out of co-optimization because market software is complex and delicate and adding features is always risky and because these system operators do not believe there are responsive loads or storage technologies that are seriously capable of supplying

[15] This is actually not true in all cases. Emissions-limited generation and responsive load, for example, may be able to supply spinning reserve but be unable to supply replacement reserve or energy due to the response duration.

ancillary services. Fortunately, Beacon Power, a developer of flywheel storage systems for supplying regulation, is having success in getting rules changed both with the ISOs and with FERC.

REGULATION AND CONTINGENCY RESERVE MARKET PRICES

NYISO hourly ancillary service market price data have been available since October 2001. Monthly averages of hourly prices are shown in Figure 15. Regulation prices, which include no fuel component, are in the same range as energy prices and are at times higher. Also, both energy and regulation prices are volatile, even on a monthly average basis. Regulation price also tends to track energy price because of the lost opportunity cost. Regulation is always the most expensive ancillary service, followed by 10-minute spinning reserve, 10-minute non-synchronous reserve, and 30-minute operating reserve.

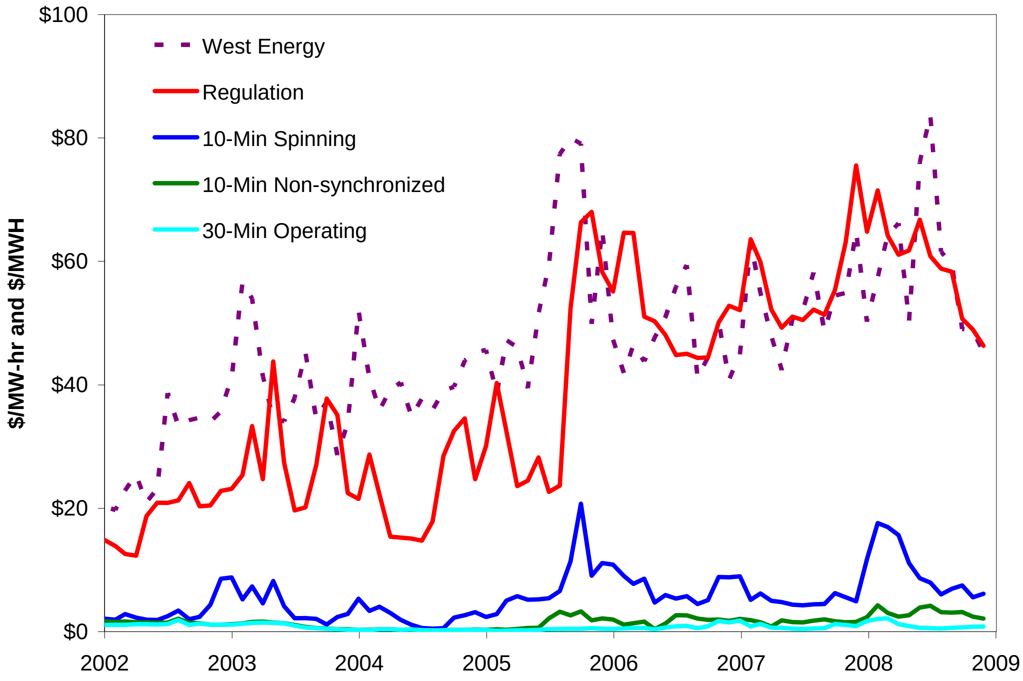

Figure 15. NYISO monthly average regulation prices are typically (but not always) somewhat lower than energy prices, and the contingency reserves are always significantly cheaper.

Figure 16 provides an average daily view of ancillary service prices from 2008. Here, the typical daily price patterns can be seen. Contingency reserve prices are typically at or near zero overnight when there is significant generating capacity that is backed down and available to ramp up. Conventional thermal plants that cannot cycle off overnight drive the price of spinning reserve down. Fast start plants keep the price of non-synchronous and 30-minute operating reserve near zero overnight. Regulation prices are higher during the morning and evening ramps in New York when the ISO is buying more of regulation. Table 4 provides a numerical comparison of the average annual prices for each service in New York and in other regions.

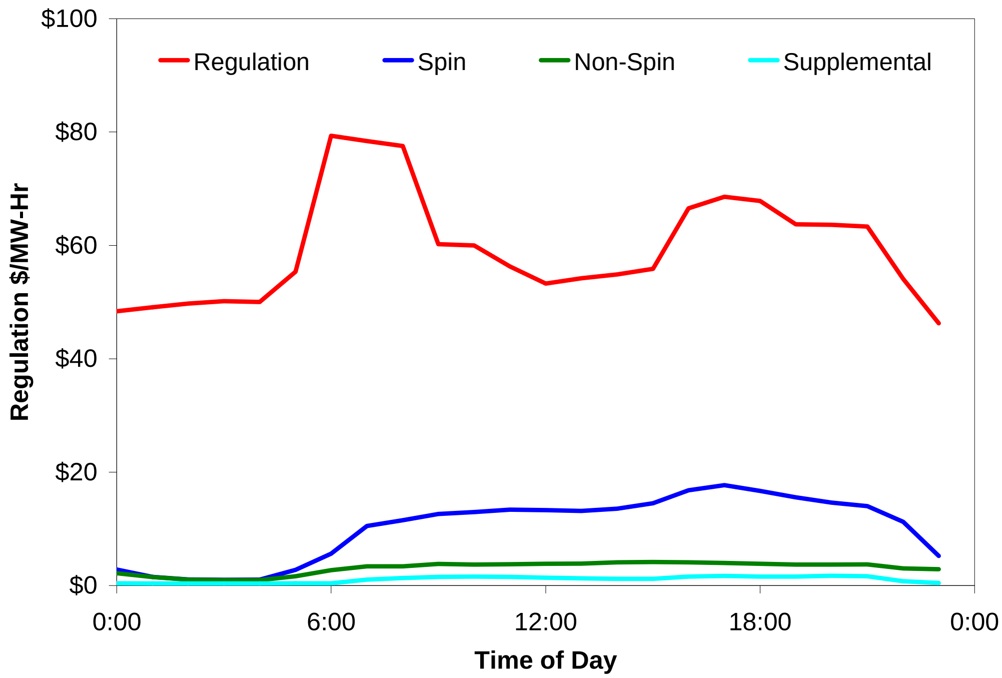

Figure 16. 2008 average hourly ancillary service prices show a consistent pattern.

REACTIVE POWER AND VOLTAGE SUPPORT COMPENSATION

While dynamic reactive power is vital for power system reliability, market-based reactive power compensation mechanisms are not yet well established or consistent. Markets are difficult to develop for reactive power because of the locational constraints. Because reactive power cannot be moved over great distances, there are usually too few generators within a given reactive power area to create a competitive market. It is enlightening to examine how generators are compensated (or not) for the provision of dynamic reactive power, though it is not completely clear if a load with reactive power capability could be compensated in the same way. Clearly, there is no incentive for a load to provide uncompensated dynamic reactive support.

Table 4. Annual average and maximum ancillary service prices from four markets for seven years

	2002	2003	2004	2005	2006	2007	2008
	Annual **Average** and Maximum $/MW-hr						
New York East							
Regulation	**18.6**	**28.3**	**22.6**	**39.6**	**55.7**	**56.3**	**59.5**
	99	195	99	250	250	300	300
Spin	**3.0**	**4.3**	**2.4**	**7.6**	**8.4**	**6.8**	**10.1**
	150	55	44	64	171	53	68
Non-	**1.5**	**1.0**	**0.3**	**1.5**	**2.3**	**2.7**	**3.1**
Synchronous	45	3	3	64	171	12	59
30 Minute	**1.2**	**1.0**	**0.3**	**0.4**	**0.6**	**0.9**	**1.1**
Operating	45	3	3	4	31	9	4
New York West							
Regulation	**18.6**	**28.3**	**22.6**	**39.6**	**55.7**	**56.3**	**59.5**
	99	195	99	250	250	300	300
Spin	**2.8**	**4.2**	**2.4**	**4.9**	**6.0**	**5.4**	**6.2**
	150	55	44	50	45	53	60
Non –	**1.4**	**1.0**	**0.3**	**0.6**	**0.9**	**1.6**	**1.7**
Synchronous	45	3	3	13	38	12	10
30 Minute	**1.2**	**1.0**	**0.3**	**0.4**	**0.6**	**0.9**	**1.1**
Operating	45	3	3	4	31	9	4
New England							
Regulation			**54.64**	**30.22**	**22.26**	**12.65**	**13.75**
(+"mileage")			344	561	100	100	100
Spin					**0.27**	**0.41**	**1.67**
					72	179	716
10 Minute					**0.13**	**0.34**	**1.21**
					50	154	716
30 Minute					**0.01**	**0.09**	**0.06**
					16	100	76
California							
Regulation	**26.9**	**35.5**	**28.7**	**35.2**	**38.5**	**26.1**	**33.4**
up+down	111	164	166	188	399	421	618
Spin	**4.3**	**6.4**	**7.9**	**9.9**	**8.4**	**4.5**	**6.0**
	250	92	125	110	225	400	400
Non-Spin	**1.8**	**3.6**	**4.7**	**3.2**	**2.5**	**2.8**	**1.3**
	92	92	129	125	110	400	399
Replacement	**0.90**	**2.9**	**2.5**	**1.9**	**1.5**	**2.0**	**1.4**
	80	55	90	36	70	175	244
ERCOT							
Regulation		**16.9**	**22.6**	**38.6**	**25.2**	**21.4**	**43.1**
up+down		177	156	1451	351	322	534
Responsive		**7.3**	**8.3**	**16.6**	**14.6**	**12.6**	**27.2**
		150	51	731	351	100	2000
Non-Spin		**3.2**	**1.9**	**6.1**	**4.2**	**3.0**	**4.4**
		249	400	510	125	180	2000

There are a range of non-market alternatives within FERC's guidelines to provide compensation to generators for reactive power capability. One approach provides no compensation to any generator and simply requires response within the (typically) ±0.95 power factor range. Another approach compensates all generators based on their revenue requirements for reactive power capability as a capacity payment (the AEP Methodology). A third approach has the system operator develop locational prices for reactive power supply based on typical cost estimates. A fourth approach periodically develops location-based prices for dynamic reactive power supply based on generator and transmission system offers (a near-market solution). There are common features for essentially all compensation systems:

- Generators are required to provide reactive response within the power factor range (typically ±0.95) if the generator is operating.
- The system operator cannot order a generator to operate simply to provide reactive support (though the system operator could ask for and pay for response)
- Generators are compensated for lost opportunities if a system operator requires an operating generator to reduce real power production to increase reactive power production.
- Cost-based reliability-must-run (RMR) contracts are used in locations where a specific generator is required to run to maintain system reliability.

For the NYISO, a resource must provide the reactive power subject to the operating conditions of the resource. A resource that wants to participate in the voltage support services also must have a functioning automatic voltage regulator (AVR) and demonstrate through a test the reactive power capability of the resource. The test validates the working order of the AVR and the ability of the resource to automatically supply reactive power and control the voltage.

Generators incur capital and operating costs for supplying reactive power. The initial capital cost of the generating plant is higher because of the required larger generator, larger step-up transformer (GSU), and additional equipment (exciter, voltage regulator, etc.) that are required to generate and control reactive power. Losses in the generator stator, rotor, and step-up transformer are incurred as operating costs. Maintenance is required for the exciter and voltage regulator. The AEP Methodology is one favored by FERC for calculating the revenue requirements associated with reactive power production. It is based on the plant capability, either nameplate or maximum obtainable.

- Reactive cost of the generator/exciter
  - =(Generator + exciter cost) * ($MVAR^2/MVA^2$)
- Reactive cost of the generator step-up transformer
  - =(GSU cost) * ($MVAR^2/MVA^2$)
- Reactive cost of the accessory electric equipment

    - =(Accessory elec. equip. cost) * (Generator|exciter auxiliary load)/(Total plant auxiliary load)
- Reactive cost of the remainder of the plant
    - =(Cost of remainder of plant) * (exciter MW/Generator MW) * (Max MVAR/Nameplate MVAR)
- Total reactive cost = Sum of the above components

Some ISOs, regional transmission organizations (RTOs) and transmission owners (CAISO, for example) provide no compensation for the supply of reactive power within a designated range.[16] Reactive supply is required by all generators as a condition of interconnecting.[17] Others (MISO for example) do compensate generators [both affiliates of vertically integrated utilities and independent power producers (IPPs)] for providing reactive power within the designated range. The institutional arrangement provides compensation using a cost-based schedule set in advance, usually a payment equal to the generation owner's monthly revenue requirement. In exchange, the generators must be under the control of the control area operator and be operated as dispatched to produce or absorb reactive power. When there is a reduction in real power output due to a request for reactive power production, the RTO will provide an additional payment to compensate the generator for the lost opportunity of delivering real power into the network. Cost-based compensation to generators for providing reactive power supply is regulated by FERC, and all ISOs/RTOs must provide a Schedule 2 tariff for reactive supply and voltage control as part of their open access transmission tariff. Examples of compensation arrangements from a number of ISOs and RTOs are presented in Table 5.

NYISO provides a $39.10/Kvar per year payment to all generators for their demonstrated reactive power capability. The payment is not restricted to locations with a specific dynamic reactive power need or denied to generators that do not operate frequently. Dynamic reactive compensation is not currently provided to any loads. No loads receive compensation for dynamic reactive capability and loads have offered dynamic reactive capability to the NYISO so the issue of loads qualifying for compensation has not come up. Technical differences between loads and generators (loads typically interconnect at lower voltages than generators) would need to be studied to determine if load compensation was appropriate. The estimated cost of providing Voltage Support Services for the NYISO in 2009 is expected to reach $62.6M, and the NYISO estimates the energy withdrawals to be 177,325,580 MWh. The resulting tariff for supplying reactive power is $0.34/MWh.

[16] CAISO does provide cost-based compensation to RMR generators for reactive support.
[17] There can be exceptions for generators that are not physically or contractually capable of providing reactive support. Wind generators get a unique "needs–based" exception from FERC.

Transmission System Reactive Power Alternatives

Generators are not the only source of dynamic reactive power. Some transmission devices can also provide dynamic reactive support to control voltages. Static Var Compensators (SVCs) and Static Synchronous Compensators (STATCOMs) are solid-state devices specifically designed to provide fast, accurate voltage control. An SVC is essentially a capacitor in parallel with a thyristor-controlled inductor. Control is fast and accurate but the device capability drops with the square of the voltage. A STATCOM utilizes a voltage source inverter so its capability drops linearly with voltage. American Superconductor offers a Dynamic VAR reactive compensation systems (D-VAR), which is a scalable STATCOM.

Being transmission-based devices, SVCs and STATCOMs are typically treated as regulated transmission assets and given a regulated rate of return. They are installed when transmission planners identify a need. They do not compete in markets to supply the service once they are installed. It is possible that appropriately equipped loads could offer to supply dynamic reactive power to a transmission owner if they were in the proper location, if they had sufficient capacity, and if the transmission owner was planning on installing a STATCOM or SVC. Payment to the loads would presumably be based on the avoided cost of the STATCOM or SVC. Figure 17presents comparative costs and sizes for various reactive power supply equipment. Note that capacitors (and reactors) are very low cost and therefore difficult to compete with if dynamic reactive power is not required.

Table 5. Regional comparison of ISO/RTO arrangements for reactive power compensation

Region	Method of Compensating Generators for Reactive Power Supply	Provisions for Testing/Confirming Reactive Power Capability of Generators & Other Facilities	Required Power Factor Capability Range for Generators (leading/lagging)	Approximate Annual Payment to Generator
NYISO	Capacity	Capability test once a years	0.95/0.90	$3,919/MVar
PJM	Payment equal to revenue requirement approved by FERC	Capability test every 5 years	0.95/0.90	$2,430/MVar
CAISO	No compensation for operating within power factor range	Tests are not normally run unless ISO detects a problem	0.95/0.90	None
ISO-NE	Capacity	Capability test every 5 years	0.95/0.90	$1050/MVar
SPP	Pass through of revenues collected by control area operators	Control area operators negotiate with generators	Not available	Not available
MISO	Payment equal to revenue requirement approved by FERC	Control area operators negotiate with generators	0.95/0.95	Generator revenues are aggregated by pricing zone

ERCOT	No capacity payment	Capability test every 2 years	0.95/0.95	Paid the avoided cost of DVAR or equivalent equipment

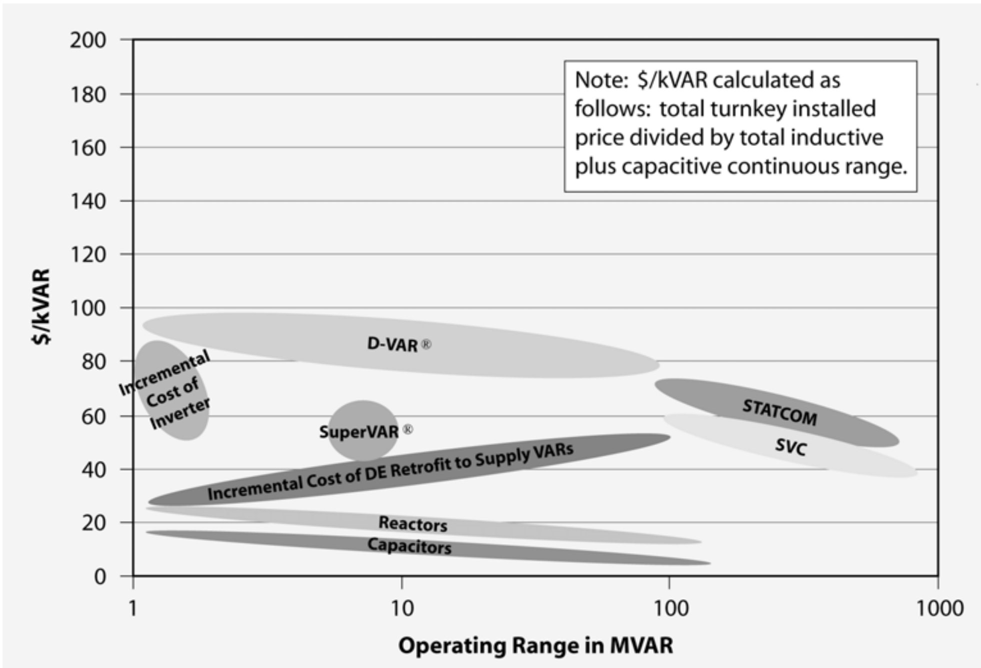

Figure 17. Relative prices of various transmission-based reactive power technologies.

5. MODELING LOAD RESPONSE WITH REAL-TIME ENERGY AND ANCILLARY SERVICE PRICES

Much can be learned by examining ancillary service and energy market clearing price data. Average regulation price data suggest that New York provides the best opportunity followed by Texas, California, and finally New England. Simple statistical analysis will not provide insights into the tradeoffs among the various response options.

The optimal selection of how a responsive load should best schedule energy use and ancillary service production while meeting its primary mission depends on the response capabilities and energy requirements of the load. Optimal response will change from hour to hour as the cost of energy and ancillary services market prices vary. An optimization model can be developed to examine how a responsive load might best respond to changing price signals.

A time series model can be developed that simulates the operation of specific responsive loads operating within the NYISO energy and ancillary services market structure. The model must reflect the load response capabilities and limitations. Communications and control capabilities must be modeled as well. Of critical importance are the load time constants and effective storage capability. Residential air conditioning loads, for example, can respond immediately but can only sustain full response for 10 to 30 minutes. Longer durations require cycling and reduced response. Energy storage is in the form of thermal energy of the conditioned space. An aluminum smelter can provide regulation and contingency response sustained for many hours but is not suited to sustained co-optimization into the energy market and thus may have to reduce its MW capability offering. Energy storage is in thermal energy for the aluminum plant as well but in the process pot line rather than in conditioned space. Other loads may have energy stored in final or intermediate product than can be stockpiled to provide longer response capability.

The model will first determine how the load would respond to the hourly energy market alone including determining the required real-power capacity. The hourly electricity procurement cost will be determined as a base line. Modeling should cover at least one year or 8760 hours.

Modeling would then examine the profit options in the regulation and contingency reserve ancillary services markets, selecting the highest profit combination for each hour. Load real power capacity would be varied parametrically to determine how the response strategy varies and how revenues increase as load capacity is increased. Interesting statistics will include revenue from the sale of each service and patterns of load operation.

Results will be time series information showing how load operations shift from hour to hour either simply purchasing energy, selling regulation, or selling contingency reserves. Hourly costs and revenues will vary as well. Detailed time series results as well as summary statistics will be available.

6. NYISO LOAD RESPONSE MARKET SPECIFICS

The preceding sections have addressed the fundamental value demand response can provide to the power system by responding to energy prices while supplying regulation and contingency reserves. We have addressed these larger trends before examining current market rules because it makes little sense to structure a business around response that is currently profitable based on matching existing rules but that actually brings little fundamental value to the power system. Use of long strings of hourly energy and ancillary service prices from mature markets assures that the analysis is fundamentally sound. Examining data from multiple mature ISOs provides further assurance that results are robust. Further, if analysis of fundamental data indicates that specific load response has value but current market or reliability rules do not allow or reward that response, then the rules probably should be changed. This section takes a brief look at the current NYISO structure for responsive load.[18] [19]

NYISO DEMAND RESPONSE PROGRAMS

The NYISO operates four programs for demand response. Participation in the four programs requires coordination, as shown in Figure 18. A load participating in the demand side ancillary services program (DSASP) can participate in either the installed capacity / special case resources program (ICAP/SCR) or the emergency demand response program (EDRP) but not both. A load cannot participate in both the DSAP and the day-ahead demand response program (DADRP) at the same time.

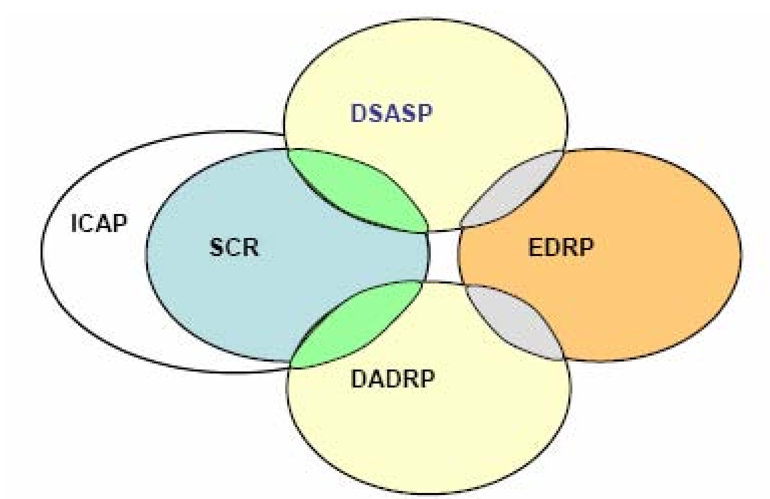

Figure 18. Participation in four NYISO demand side programs requires coordination.

[18] Donna Pratt, 2008, "Introduction to Demand Side Ancillary Services," DSAP Training Course, NYISO, November 21.
[19] NYISO Market Participants User's Guide http://www.nyiso.com/public/webdocs/documents/guides/mpug_mnl.pdf

DSASP: demand side ancillary service program – Loads selling regulation and the contingency reserves. No loads currently supply ancillary services to the NYISO but three are close to being qualified.

DADRP: day-ahead demand response program – Loads sell curtailment back to the NYISO day-ahead at energy market prices.

- Bids to curtail if called upon are submitted by 5am the day ahead.
- There are financial penalties for failure to perform.
- DSASP resources <u>*cannot*</u> participate in the DADRP program (if called on to curtail they would not be available to provide the ancillary services).

ICAP/SCR: installed capacity/special case resource (ICAP/SCR) is one of the reliability based demand response options of NYISO[20]. The special case resources (SCRs) comprise of the end-use interruptible loads, who agree to reduce electricity consumption in the case of emergencies, and the distributed generators. The rated capacity of these resources should be at least 100 kW or higher. A group of small end users may also qualify for this program following the 100 kW capacity criteria. The SCRs receive capacity payments from the NYISO for the load curtailment commitment. For 2008, the average rest–of –system weighted average price was $2.27/kW-mo with summer prices being higher than winter. The guaranteed amount of load reduction for a certain contracted period by the SCRs is referred to as the unforced capacity (UCAP) which can be freely traded in bilateral transactions. Separate entity termed as responsible interface parties (RIPs) stand as a communication hub of NYISO for the verification of SCRs and UCAP associated with the SCRs and dispatch of SCRs including other administrative duties. Following are some requirements which SCRs have to adhere to in NYISO:

- The UCAP of SCRs can only participate in the auctions managed by the NYISO
- The load-serving entity (LSE) can claim for its UCAP obligation only in an even increments of 100 kW, that is, 395 kW of UCAP would be rounded to 300 kW
- The aggregated SCRs are allowed only to fulfill the 100 kW block criteria but is not allowed to compensate for the underperformance of some SCRs by other SCRs which are over performing
- The SCRs must provide the energy for a minimum of 4 hour block unless they have been approved by NYISO for some shorter blocks due to some constraints

The NYISO provides at least 2 hour notice to the RIP when the Special Case Resource action is required and the SCR should start the curtailment of the required amount of load or the transfer of the load to the distributed generators immediately after the 2 hour duration has expired. If the SCR fails to respond to the RIP notification in order to provide the agreed load reduction or the committed maximum increase in

[20] *Installed Capacity Manual*, April 2009, NYISO. http://www.nyiso.com/public/webdocs/products/icap/icap_manual/icap_mnl.pdf, accessed on July 2009.

distributed generator output, SCR will be considered forced out while calculating the UCAP value of the SCRs. This will have financial implications to the SCRs.

The Special Case Resources receive a payment via SCR's RIP from the NYISO for every task performed if the SCR's data submitted by RIP is within the pre-specified time frame and could be verified properly. The energy payment is done based on the amount of load curtailment for the specified time period measured in terms of energy supplied during each hour for the entire duration of the event. Upon the request from NYISO to perform for less than the minimum of 4 hour block, SCRs are paid as if it has been in service for the entire 4 hours. Once the SCRs are verified for the duration of service as requested by NYISO, every SCR is paid the zonal real time locational based marginal pricing (LBMP) per MWh of load reduced.

EDRP: Emergency Demand Response Program – Loads are paid the greater of $500/MWH or the LBMP when called on by NYISO to voluntarily curtail.

- There is no penalty for failing to comply with a EDRP request (other than not getting the $500/MWH)
- DSASP resources *can* participate in EDRP. EDRP and ICAP/SRC are mutually exclusive.

Aggregators account for 92% of the participating resources and 70% of the registered MW in the ICAP/SCR program. Total demand response is growing. There is a consistent trend both in the number of participants and in the total MW response in moving from the EDRP program to the ICAP/SCR program as shown in Figure 18. The ICAP/SCR strike price was $500/MWH for the 1,744MW of participating load in 2008. The DADRP program has remained relatively stable with 50 participants offering 331 MW through 11 curtailment service providers in 2008. Compensation averaged $104/MWH for scheduled energy. Only 7,727 MWH of DADRP were scheduled in 2008 so, DADRP is used relatively sparingly.[21]

Demand side resources (DSASP) are not paid for energy when it is actually scheduled. The theory is that they will not pay for energy they do not consume themselves. This prevents responsive loads with conventional power supply arrangements from capturing the full value of their response. Loads do put in a minimum $75/MWH energy bid that reflects their willingness to be curtailed.

A resource providing regulation or spinning reserve may have to buy out of its energy obligation in real time as it follows ISO dispatch instructions. The day-ahead margin assurance payment (DAMAP) is designed to protect a resource's day-ahead margins. For generators, the DAMAP is equal to the cost to buy out energy in real time less the day-ahead payment plus the lost profit.

[21] Donna Pratt, 2009, "2008 Year-End FERC Report on NYISO's Demand Response Programs", Price Responsive Load Working Group, NYISO, January 29

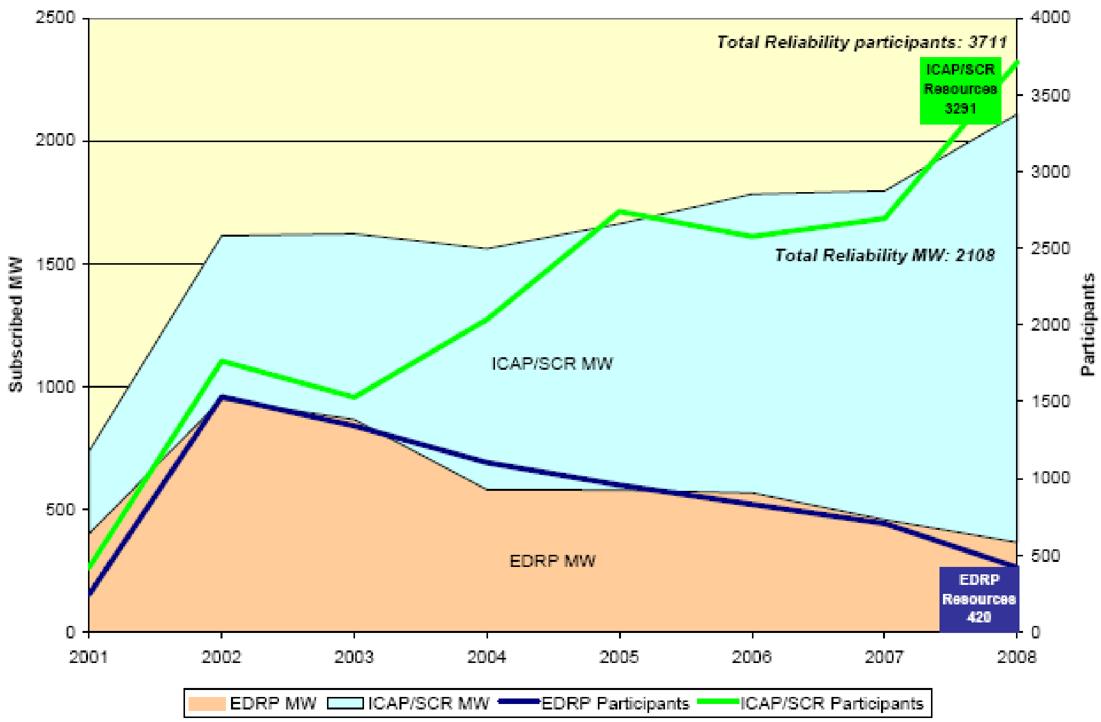

Figure 19. Total demand response is growing. There is also a trend to increase ICAP/SCR and decrease EDRP participation.

Demand side resources are not evaluated as energy providers in the day-ahead market. This does provide some partial protection from co-optimization. The actual bidding requires supplying information that describes the resource as though it was a generator with minimum and maximum load, emergency upper operating limit, 10 minute ramp rate, minimum run time, etc. Unlike generators, loads bidding regulation or spinning reserve must have a minimum generation cost of $0/MWH. While this may be appropriate for some loads, it fails to reflect load opportunity costs or other operating expenses a load incurs while adjusting its process to accommodate the power system. Unlike generators, loads bidding regulation or spinning reserve must have a start up cost of $0. The bid curve can be up to 11 blocks of energy prices. The current floor is $75/MWH.

NYISO schedules every 5 minutes in real-time with 3 possibilities every 5 minutes: not used, converted to energy, and held as reserve. For resources supplying contingency reserves, there is no guaranteed tie between being used and a reserve pickup event but deployment is much more likely during an event.

PERFORMANCE

A demand side resource has a performance index that is calculated for each interval of its real-time demand reduction schedule. The index is used in the DAMAP calculation. If a resource does not perform adequately, these payments made to the demand side resources may be reduced. Knowledge of the calculation methods for the performance index can provide insight on key requirements for delivering the acceptable demand response.

Reserve Performance Index

The reserve performance index identifies the performance of the resource in providing operation reserve. The scale of the performance index is from 0 to 1. A value of 1 is given to resources that perform the task effectively, while 0 is assigned to resources that fail to meet the signaled response.

The performance index is a composite of the actual demand reduction (ADR) and the ramped scheduled reduction (RSR) over an interval. A resource that actually provides more than the required response is not penalized. The equation is as follows:

$$PI_i = Min\left[(ADR_i / RSR_i + 0.10), 1\right]$$

An example system AGC and resource response are provided in Figure 20. In this case, the resource missed the mark during the first 20 minutes, but was able to follow the AGC during the last 10 minutes. The performance index for each 5 minute interval is shown in

Table 6. This index can be utilized to penalize resources financially when performance is not acceptable.

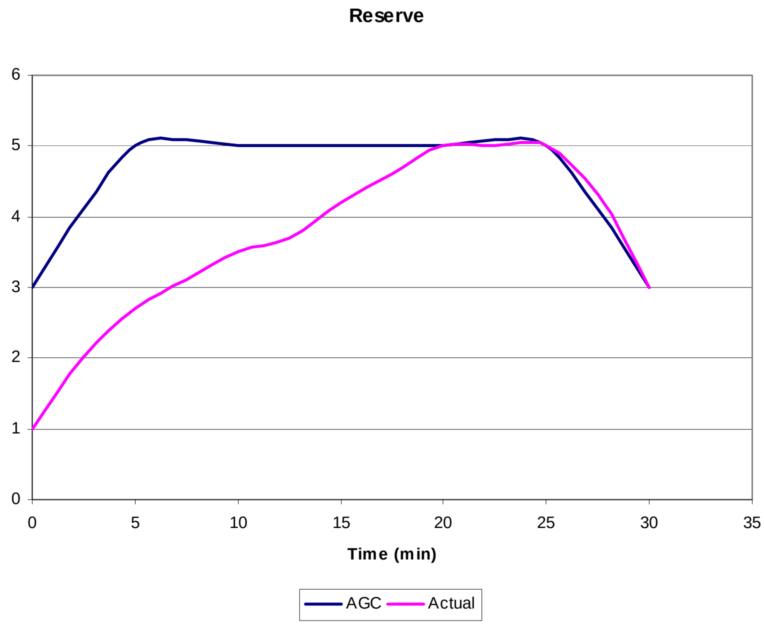

Figure 20. Reserve example AGC and resource response.

Table 6. Calculation of performance index during each 5 minute interval.

	PI	
Interval 1	0.59	= Min[(2 / 4.1)+0.1, 1] = Min [0.49+0.1,1] = Min[0.59,1]
Interval 2	0.71	= Min[(3.1 / 5.1)+0.1, 1] = Min [0.61+0.1,1] = Min[0.71,1]
Interval 3	0.84	= Min[(3.7 / 5)+0.1, 1] = Min [0.74+0.1,1] = Min[0.84,1]
Interval 4	1	= Min[(5/ 5)+0.1, 1] = Min [1+0.1,1] = Min[1.1,1]
Interval 5	1	= Min[(4.3 / 4.1)+0.1, 1] = Min [1.1+0.1,1] = Min[1.2,1]

Regulation Performance Index

The regulation performance index tracks how well a regulation supplier responds to the control signals that are issued every six seconds. The performance index is calculated through examination of the positive control error (over-generation) and negative control error (under-generation) over a 30 second interval. OG and UG cannot be below zero.

$$OG = MW_{meas,p} - BP^{+}{}_{AGC30} \text{ (over generation)}$$

$$UG = BP^{-}{}_{AGC30} - MW_{meas,p} \text{ (under generation)}$$

These values are accumulated for each 30 second interval in the RTD period to create the positive and negative control errors.

$$PCE_i = \sum OG$$

$$NCE_i = \sum UG$$

The regulation performance index also utilizes the amount of regulation the provider output could change during the interval or the unit regulation margin. In this equation, RR is the regulation ramp rate (MW/min) and s_i is the number of seconds in the RTD interval.

$$URM_i = RR \times [s_i / 60]$$

The final equation for the regulation performance index is as follows:

$$PI_i = \left[\frac{URM_i - (PCE_i + NCE_i)}{URM_i} + 0.10 \right] \times \left[\frac{\mathrm{Re}\, gPeriod_i}{s_i} \right]$$

An example AGC signal and load response waveforms for regulation are provided in Figure 21 and the data is shown in Table 7. Each 6 second interval is taken and averaged into a 30 second interval for the MW measured. This is compared to the maximum and minimum basepoints. In this case, the minimum and maximum basepoints were not exceeded; therefore the OG and UG values are zero. The calculations are shown in Table 8.

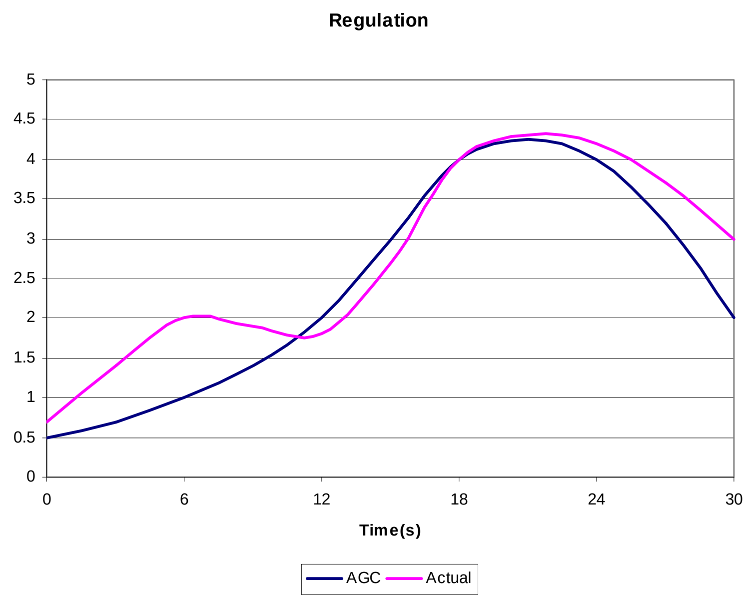

Figure 21. Example AGC and response curves.

Table 7. Example AGC and load response figures

Interval	1	2	3	4	5
AGC	.5	1	2	4	4
Actual	.7	2	1.8	4	4.2

Table 8. Calculation of the PCE and NCE over a single interval

	Settlement	
$MW_{meas,p}$	2.5	= average of actual response for intervals 1 through 5 intervals
BP^+	4.0	= maximum AGC basepoint for intervals 1 through 5
BP^-	0.5	= minimum AGC basepoint for intervals 1 through 5
OG	-1.5	$= MW_{meas,p} - BP^+ = 2.5 - 4.0$
UG	-2.0	$= BP^- - MW_{meas,p} = 0.5 - 2.5$
PCE_i	0	Since OG < 0, OG = 0
NCE_i	0	Since UG < 0, UG = 0

ECONOMIC EXAMPLE

The NYISO utilizes particular economic strategies or rules to evaluate the financial reward and cost to loads participating in DADRP and DSASP. Examining these rules is conceptually important to understand the risks and rewards towards a demand side resource.

DADRP

A load participating in DADRP is charged, as if under normal conditions, the price agreed upon for the MW utilized. However, if the load is called upon to curtail, the load is both credited the amount of load curtailed and paid a performance payment.

The market rules also require that the load be charged a balancing fee for not remaining during the schedule. To ensure that the load is not cheated, this fee is reimbursed to the LSE or demand response program (DRP). The load can also receive further compensation if the curtailment initiation cost and lost opportunity cost do not exceed the performance payment.

For example, a load that is approximately 20MW with the capability to curtail 25% or 5MW can either be rewarded or punished based on performance. In one case, the load schedules to curtail 5MW for 8 hours with a day-ahead locational marginal price (LMP) Price Cap of $150/MWh and has a curtailment initiation cost of $3500 for the 8 hours. This results in a $9500 total curtailment bid. The calculations and assumptions are shown on Table 9.

If the market day-ahead and real-time price are $200/MWh and $250/MWh and the load is called on to curtail 5MWh for the 8 hour, the demand side resource has the potential to reduce the cost by more than half while only shedding 25%.

Initially, the LSE/DRP must pay $32,000 for the load but is compensated for the reduced load with 2 payments of $8,000 (one for adjustment of the load and the other as a performance payment.) Since the curtailment bid exceeded the compensation payment in this example, an additional payment, or the difference, is also paid to the LSE/DRP. Finally a load balancing credit and debit are added to the total to create a final payment to the load of $14,500 as shown on Table 10.

Table 9. Example assumptions

	Day Ahead	Real-Time	
LMP($/MWh)	200	250	Assumed Example
Fixed Load (MW)	20	20	Net Load + Reduction
Load Reduction (MW)	5	5	Actual Reduction
Total DAM Load	15	15	Real Time Net Load
Shutdown Duration (hrs)	8	8	Assumed
Hourly Curtail Bid	$150		Price Cap
Curtail Initial Cost	$3,500		
Total Bid Cost	$9,500		= 8 hours x $150/MWh x 5MW

Table 10. Settlement process for successful curtailment

	Settlement	
LSE DAM Purchase Obligation	($32,000)	= 8 hours x $200/MWh x 20MW
LSE DAM Credit	$8,000	= 8 hours x $200/MWh x 5MW
Payment for Performance	$8,000	= 8 hours x $200/MWh x 5MW
Nonperformance Penalty	$0	The load delivered as promised.
Hourly Bid Curtailment Cost Guarantee Payment	$1500	= $9500 - $8000
LSE Load Balance Credit	$10,000	= 8 hours x $250/MWh x 5MW
LSE Load Balance Debit	($10,000)	= 8 hours x $250/MWh x 5MW
Total Received (paid) by LSE	($14,500)	= $32,000 - $8,000 - $8,000 - $1500 - $10,000 + $10,000

If the load fails to curtail when required, a penalty fee is attached, and the LSE does not receive a performance payment as shown in Table 11. The penalty rate is the higher of the day-ahead and real-time price. as a result, the LSE must actually pay more than the original designated price of the energy for the load. It is important to ensure that the load reaches the required response when participating in DADRP to establish a financial opportunity.

Table 11. Settlement process for failed curtailment

	Settlement	
LSE DAM Purchase Obligation	($32,000)	= 8 hours x $200/MWh x 20MW
LSE DAM Credit	$8,000	= 8 hours x $200/MWh x 5MW
Payment for Performance	$0	The load does not curtail
Nonperformance Penalty	($10,000)	= 8 hours x $250/MWh x 5MW
Hourly Bid Curtailment Cost Guarantee Payment	$0	The load does not curtail
LSE Load Balance Credit	$0	The load does not curtail
LSE Load Balance Debit	$0	The load does not curtail
Total Received (paid) by LSE	($34,000)	= $32,000 - $8,000 + $10,000 - $10,000 + $10,000

METERING AND COMMUNICATIONS

The communications and metering requirements for selling ancillary services do not appear to be designed to support large aggregations of small responding loads (such as home water heaters). This could be interpreted to mean that small loads cannot participate. Alternatively, it could mean that someone proposing to use an aggregation of small loads needs to discuss the requirements with NYISO and attempt to establish appropriate requirements.

Metering Configurations

For participation in the DSASP, the NYISO has several specific requirements and metering configurations that can be utilized. The local transmission owner should be contacted for specific device and installation requirements, but Figure 22 shows several standard configurations that can be utilized to provide regulation and synchronous reserves. The first configuration has the total load connected to a RTU/Instantaneous Meter through a utility revenue meter. The RTU communicates the necessary information to the transmission owner and ISO and should be capable of scanning output data every six seconds. The second and third configurations employ a meter data service provider (MDSP) revenue-grade meter in place of the utility revenue meter. The utility revenue meter is still utilized, but does not supply the RTU/instantaneous meter. The difference in the second and third configurations is the sub metering of the loads. Sub metering of the loads requires a sub meter totalizer to submit a signal to the MDSP revenue-grade meter. A list of the revenue meters can be found at http://www.dps.state.ny.us/approved_meter_list.PDF.

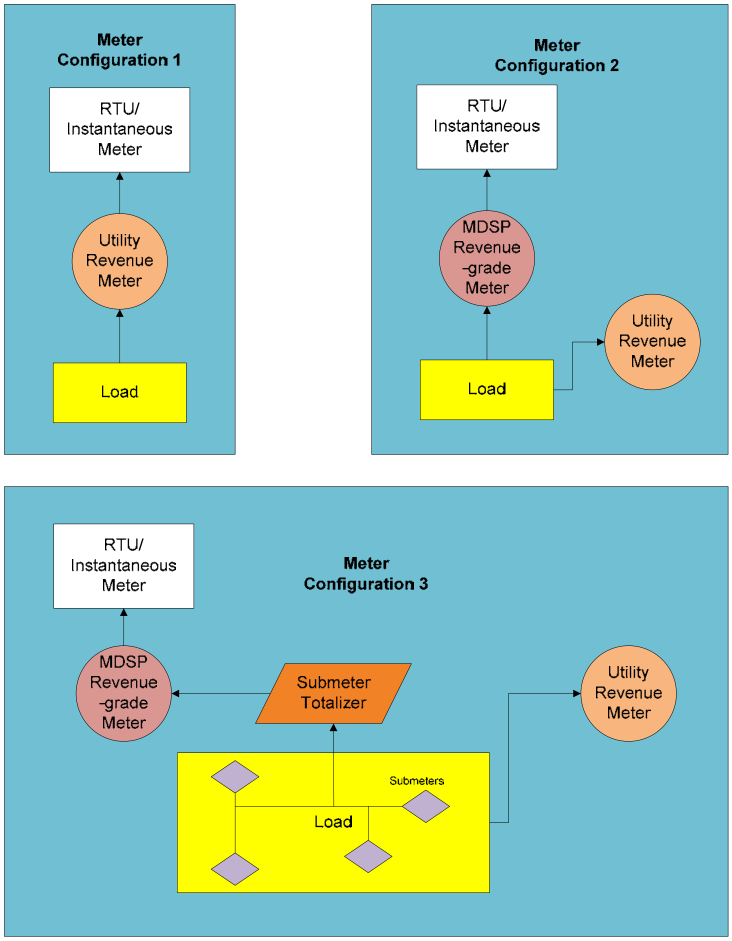

Figure 22. Metering configurations for regulation and spin supported by NYISO.

Communication Configurations

Communication requirements also exist for involvement in DSASP. Figure 23 has several examples of configurations that can be adopted. The chosen configuration must be capable of receiving automatic control signals on a 6 second periodicity for regulation and a 5 minute periodicity for operating reserves and providing telemetered output data every 6 seconds either through the transmission operator or directly through the NYISO.

The parameters transmitted every 6 seconds include, a regulation flag that specifies when a regulation schedule is active, a base load, and the required response. The regulation flag will read either as a zero or one, zero for no regulation service, or a one for regulation scheduled.

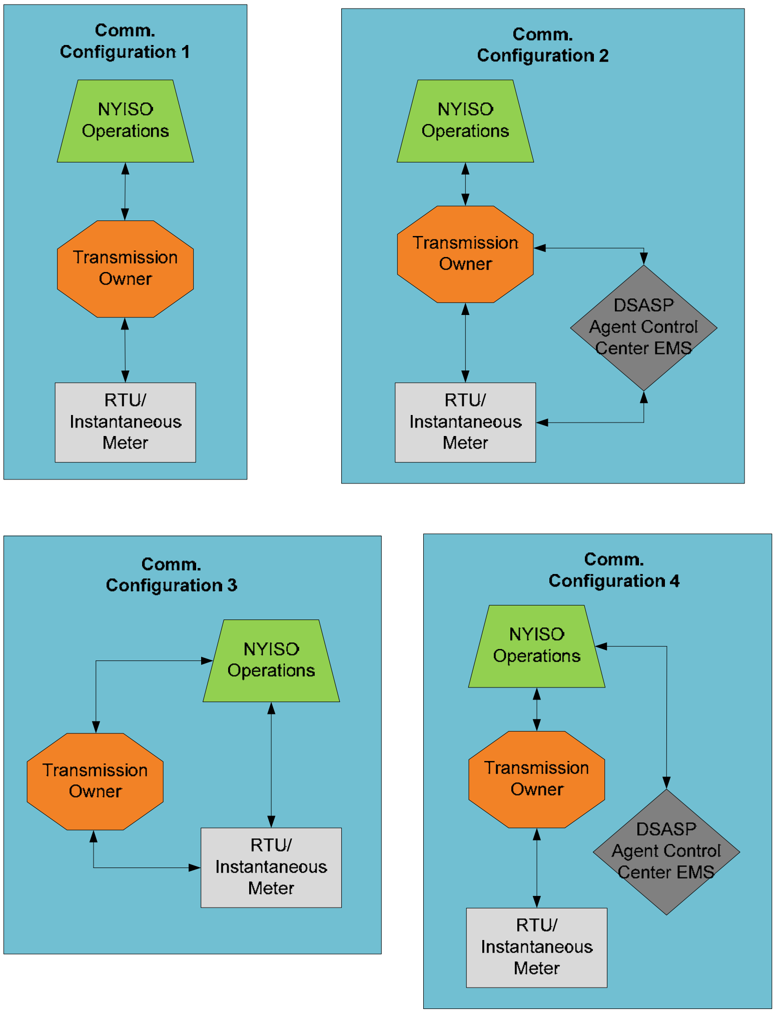

Figure 23. Communication configurations supported by NYISO.

PRE-QUALIFICATION TESTING

Resources wishing to sell ancillary services must pass pre-qualification testing. For regulation this involves successfully supplying 100 hours of regulation service within 1 month. For contingency reserves this consists of responding to a test event within the 10 or 30 minutes (depending on the reserve being supplied) with 1 or 3 minute tolerance and within 2% of the specified response mount. In supplying reactive power, the resource must operate at maximum leading and lagging power factor for one hour. The NYISO can conduct performance tests at any time.

7. SUB-HOURLY ENERGY MARKETS

Sub-hourly energy markets show surprising price volatility. It may be possible for responsive loads to profit by responding to this volatility though the exact mechanism has not yet been explored.

NYISO (and most other ISOs) operate a set of near real-time energy markets: day-ahead hourly, hour-ahead hourly, and 5 minute sub-hourly energy markets (ERCOT currently operates a 15-minute energy market). Prices from these three markets are typically fairly close on an annual average basis as shown in Table 12. The annual average 5-minute price ($63.31/MWH) was quite close to the annual average hour-ahead-hourly price ($64.93/MWH) and not far from the day-ahead-hourly price ($67.70/MWH).

Table 12. Annual average prices for 2008 show consistence between energy markets. Sub-hourly markets show strong volatility

ISO	Day-Ahead $/MWH	Hour-Ahead $/MWH	5-Minute $/MWH	Average Within-Hour 5-Minute Range $/MWH
NYISO	$67.70	$64.93	$63.31	$91.18
ISO-NE	$81.38	$80.76	$81.22	$24.40
CAISO		$69.78	$68.32	$59.87
ERCOT[1]			$71.69	$40.00
MISO	$49.99	$48.62	$48.71	$67.75

[1]ERCOT currently operate a 15 minute sub-hourly market rather than a 5 minute market.

The right hand column in Table 12 shows the average of the 5 minute price range each hour: $91.18/MWH. That is, the average difference between the highest and lowest of the 12 5-minute prices for each hour was $91.18/MWH for all of 2008. This is quite a bit of price volatility.

Sub-hourly price volatility does not show much of a clear pattern in NY (the pattern is clearer in other ISOs) as demonstrated by Figure 24. The left hand figure shows the absolute value of the difference between the sub-hourly interval price and the average price for the hour while the right hand figure shows the signed difference. The two horizontal axes are hour-of-the day and minutes-into-the-hour respectively. The vertical axis shows the price difference itself. The lack of daily or hourly pattern, coupled with the fact that the average sub-hourly price range shown in Table 12 is greater than the systematic price pattern shown in Figure 24 indicates that the 5-minute variability in NY is similar to regulation which is highly random.

This 5 minute price volatility provides a very strong economic signal for generators with a marginal cost close to the market clearing price to move their generation output as a short-term strategy to chase the 5-minute price trend. While regulation is the most expensive ancillary service (generators and responsive loads are explicitly paid to provide regulation), sub-hourly energy markets extract load following response at little or no cost. Generators still profit from responding to sub-hourly energy price volatility, however.[22]

Sub-hourly energy markets send a strong economic signal that favors response. It is not clear if responsive loads can profit from this price volatility. Maximizing profits might require multiple coordinated energy transactions.

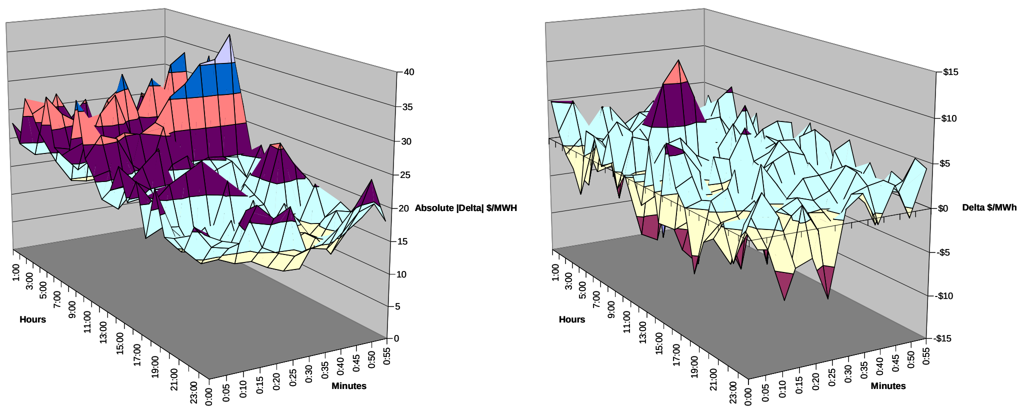

Figure 24. NYISO 5 minute prices are highly volatile.

[22] B. Kirby, M. Milligan, 2008, *Examination of Capacity and Ramping Impacts of Wind Energy on Power Systems*, NREL/TP-500-42872, July.

8. INDUSTRIAL LOAD OPPORTUNITIES IN NYISO ENERGY AND ANCILLARY SERVICES MARKETS

Industrial loads (and all loads) have significant opportunities to reduce their energy costs in New York by responding to power system needs. Both response capability and actual performance are valued and paid for. With the opening of the ancillary service markets to load participation, loads can better tailor their response to maximize both the payments they receive and the reliability benefits they provide to all customers.

Figure 18 shows the four program options loads have for providing response. The programs were described in Section 6. Here the response options will be compared to better understand the potential income offered by each as well as the response requirements and costs. Table 13 provides a comparison of response compensation and the approximate total response requirement (or current market size).[23]

DSASP: DEMAND-SIDE ANCILLARY SERVICE PROGRAM

Ancillary service response is the highest priced and potentially the most profitable response for loads to provide. Response speed and accuracy are required, increasing the communications and control requirements and limiting the types of loads that can provide the services. Selling ancillary services essentially provides a dollar-for-dollar reduction in energy price, based on the ancillary service market clearing price, for every megawatt of service the load is able to provide. Responsive loads will likely incur increased operating costs when they provide ancillary services, but the ancillary service price will more than compensate for any increased costs for some responsive loads.

Regulation

Regulation is the by far highest priced and most demanding ancillary service: $59.46/MW-hr on average. Regulation requires essentially continuous minute-to-minute up and down following of the NYISO AGC signal, so most loads will not be able to provide it. Loads that can provide regulation must have excess capacity; they operate at a lower average power level and swing up to their peak consumption. The regulation price is not constant throughout the day but, unlike contingency reserves, it does stay high at night (Figure 16).

A very flexible industrial load with a 1 MW average energy requirement, for example, could sell regulation and reduce its energy cost by about $60/MWH on average. To do so the load would have to be capable of swinging between zero and 2 MW while still averaging 1 MW in order to continue performing its basic function (making aluminum, for example). That is a dramatic energy cost reduction that requires exceptional flexibility as well as excess process capacity. A less flexible load might average 10 MW while

---

[23] Prices are shown for western New York. Prices for some services such as the contingency reserves are somewhat higher on Long Island and in New York City.

swinging between 9 MW and 11 MW in order to supply 1 MW of regulation. The load would still receive an average $60/MW-hr, which would now offset the cost of 10 MW of energy resulting in an energy price reduction of ~$6/MWH for the entire load. Clearly, there is an incentive to supply as much response as possible.

Table 13. NYISO response compensation (2008) is higher for faster, more dependable performance

Response	Compensation	Market Size
DSASP (west)		
Regulation	$59.46/MW-hr annual average	150-275 MW
Spinning Reserve	$6.19/MW-hr annual average	600 MW
10-Minute non-sync	$1.72/MW-hr annual average	600 MW
30-minute	$1.06/MW-hr annual average	600 MW
DADRP	No capacity payment + $104/MWH average energy market price during relatively rare events	311 MW
ICAP/SCR	$3.11/MW-hr annual average + $500/MWH average when called	1,744 MW
EDRP	No capacity payment + $500/MWH or LBMP when called	600 MW

Contingency Reserves

Loads and generators can be paid to stand ready to respond in the event a generator or transmission line fails. Faster response capability receives higher compensation with 10–min spinning reserve paid $6.19/MW-hr on average while 30-min operating reserve receives only $1.06/MW-hr. Prices are typically higher from 6am to 10pm and are often zero overnight. Though the load must stand ready to respond, actual response is required only if a power system event occurs. While the contingency reserve prices are well below those paid for regulation, the load does not need to have capability to increase consumption; only curtailment is required. Unlike the other demand response programs, curtailment duration for contingency reserve events is typically limited to a fraction of an hour though 2 hour curtailments are possible. For some loads, the curtailment duration is critical and these loads must avoid being co-optimized into the energy market.

Accounting for the energy not consumed during contingency response is handled separately from the payment for standing ready to respond. While it is important to account for the energy accurately, the saved energy is typically not a major economic factor in deciding which type of response to provide. In fact, the lost production consequences for the load are typically more important economically than the saved energy.

The higher price ($6.19/MW-hr annual average) and typically shorter response duration favors supplying spinning reserve as opposed to non-synchronized or 30–min operating reserves. The load must be capable of beginning to respond immediately and response to frequency decline must be automatic. Payment is made for each hour the load stands ready to respond, not just when response is required. While payment is lower than for regulation, response is much simpler as well: it is only necessary to rapidly turn off the load. A greater range of industrial, commercial, and even residential loads are potentially capable of supplying spinning reserve and loads that can provide both can typically provide more contingency reserves than regulation.

DADRP: DAY-AHEAD DEMAND RESPONSE PROGRAM

The day-ahead demand response program allows loads to sell curtailment back to the NYISO day-ahead at energy market prices. A load cannot offer the same response into both the DADRP and the DSASP since the response capability, if selected to provide energy curtailment, would be unavailable to provide ancillary service response. Curtailments are typically multiple hours when they occur, but the program is used relatively sparingly. Compensation averaged $104/MWH in 2008, but the number of hours was limited.

ICAP/SCR: INSTALLED CAPACITY/SPECIAL CASE RESOURCES

Loads are allowed to offer (and get paid for) UCAP in the NYISO semiannual auction. The average price for 2008 was $2.27/kW-month, which is the equivalent to $3.11/MW-hr. Loads also get paid when they are actually required to curtail. Curtailments, when required, typically last for several hours. Loads can provide both ICAP/SCR and ancillary services. ICAP/SCR response takes precedence over ancillary service supply if both are required simultaneously.

EDRP: EMERGENCY DEMAND RESPONSE PROGRAM

The EDRP program provides no capacity payment and imposes no response obligation. Loads can respond when they want if the price is attractive (the greater of $500/MWH or the LBMP when called).

VALUING RESPONSE AND MAXIMIZING PROFITS

Load can reduce their electricity costs by responding to power system reliability needs. Table 13 compares the average annual revenue a load can receive by providing different types of response. Prices vary from season to season and, for ancillary services, from hour to hour. Compensation for some response can be very high. EDRP, for example provides at least $500/MWH but only for a few hours per year. Regulation, on the other hand, pays nearly $60/MW-hr (on average) 8760 hours per year. Response speed, duration, and warning differ as do the capabilities of individual loads to provide the specific responses. Maximizing profits requires analyzing both the cost and revenue for each specific load. Response must also be integrated with the load’s energy procurement and use strategy. The most profitable response will likely vary from hour to hour and season to season. Modeling can be a great help in optimizing performance.

9. CONCLUSIONS: ANCILLARY SERVICE STRATEGIES FOR RESPONSIVE LOADS

Sale of regulation and spinning reserve may be profitable for responsive loads. Preliminary examination of ancillary service prices in New York, Texas, California, and New England indicate that regulation prices are particularly attractive, and have been for years. Study of the hourly energy and ancillary service market prices provides a sound basis for determining the genuine value of response to the power system and the revenue responsive loads can likely capture.

As a general rule, the ancillary service and energy markets are reasonably mature and well behaved. Ancillary service prices behave as theory predicts: ancillary service prices are driven by generator opportunity costs in the energy markets. This provides further confidence that the value of response is both real and long term providing a reasonable basis for long term decisions.

Because the interaction of load requirements, energy prices, and ancillary service prices are complex, it would be best to model industrial load operations to determine what capabilities and operating strategies will likely be most beneficial. Modeling should be conducted for at least a year of hourly energy and ancillary service price data. Each load's capabilities and limitations need to be quantified to support this modeling effort.

Actual market and reliability rules complicate the analysis of what market participation is currently practical. It is important to understand the specific ISO rules for three reasons. First, any immediate market participation must comply with the existing rules. Second, some differences result from differences in market design and may continue in the foreseeable future. The inclusion of capacity payments in New York, for example, represents a basic market design philosophy, and responsive load probably needs to find ways to maximize the capacity payments to maximize revenue. Third, it is necessary to understand limitations that the existing rules impose on responsive loads that are not truly related to reliability requirements and that only reflect the capabilities of the historic resources that supplied the response services (conventional generators) if these rules are to be changed. It is worth noting that Beacon Power appears to be having success in getting regulation supply rules changed to reflect the energy limitations of its flywheel storage system.

Responsive loads with adjustable speed motor drives, solid state power supplies, or synchronous motors potentially can supply dynamic reactive power to control voltage and support reliability. Dynamic reactive power requirements are local and reactive support markets have not developed. Dynamic reactive capability is likely to be a capability that loads should always offer to the local transmission owner. Compensation will likely have to be negotiated each time. Compensation can be based on alternative costs from generators (opportunity costs for RMR generators) or capital costs from transmission devices (SVCs and STATCOMs). Further analysis can identify locations where dynamic reactive supply is likely to be valued.